University of Alberta

Children's Misunderstanding of the Equal Sign and Their Adherence to Addition Schemas: Detecting the Effects of Prior Arithmetic Experience and Instruction

by

Jody Lyn Sherman ©

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Psychology

Edmonton, Alberta
Fall 2007

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-33065-4
Our file Notre référence
ISBN: 978-0-494-33065-4

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Dedication

To those who arrive at work every morning, staring out over a sea of eager young faces, trying to impart all that has been entrusted to them to teach, and to the children who must navigate through this world of arbitrary symbol systems in order to *understand,* thank you for sharing your wisdom with me. I have tried to capture your strategies, perceptions, and beliefs here, with the hopes of influencing instruction for the future. Undoubtedly the picture I present is so shallow and summative that some of your individuality may appear unnoticed or unimportant. That couldn't be further from the truth. Even so, thank you.

And to my Grandma, who always wanted me to be a "doctor." I miss you.

Abstract

Elementary school children's misunderstanding of the equal sign and their failure on equivalence problems (e.g., $a + b + c = a + \_\_$) have been well documented and are associated with difficulty in algebra in higher mathematics. In a series of five studies I explored children's early ability to solve equivalence problems, the benefits of experiencing equivalence problems in less symbolic formats prior to symbolic problems, children's performance on equivalence problems over time in relation to specific interventions, and teachers' instructional practices with regard to equivalence. In Study 1, a *nonsymbolic* condition (i.e., blocks and bins rather than conventional mathematical symbols) was used to assess whether kindergarten children can solve equivalence problems. Almost 30% of the students solved more problems correctly than would be expected by chance, indicating that many children begin their formal schooling already knowing how to make two sides of an expression equivalent. However, young students seem unsuccessful in mapping this understanding to the equal sign in later years. In Studies 2 and 3, Grade 2 children solved equivalence problems presented either nonsymbolically or with just some symbols (*semi-symbolic*) in one session and typical symbolic problems in another session. Children who started with the less symbolic conditions had very high accuracy, especially on symbolic problems presented a week later. Benefits of prior experience with few or no symbols on subsequent symbolic problems were examined in more detail using a microgenetic design. In Study 4, children's changing views about the equal sign were examined over several months. Although most children's ability to solve and reconstruct equivalence problems improved over time, boundary conditions surrounding the effectiveness of the nonsymbolic and

semi-symbolic interventions, such as the potential importance of multiple-choice solution options, became apparent. In Study 5, most teachers indicated they were unaware of students' misunderstanding of the equal sign. Thus, children's view of the equal sign may improve after exposure to certain conditions, these conditions may be subject to important boundary conditions, and teachers' ability to help children map the concept of equivalence to the equal sign may be limited because teachers are unaware of children's misperceptions.

Acknowledgment

My primary thanks and appreciation go to my supervisor, Dr. Jeffrey Bisanz. His support, encouragement, and enthusiasm were instrumental in my research and career decisions, and his patience and creative energy were exactly what I needed to complete the studies reflected in this project.

I would also like to thank Dr. Elena Nicoladis, and Dr. Connie Varnhagen for serving on my supervisory committee. They provided exceptional feedback, guidance, and support throughout my graduate degree. I also extend my gratitude to Dr. Jacqueline Leighton and Dr. Martha W. Alibali for their service as external committee members and their thoughtful revision suggestions.

I would like to acknowledge members of the Mathematical Cognition Lab (Carmen Rasmussen, Rebecca Watchorn, and Katy Wyper) for their helpful comments and feedback throughout this and related projects. Ozlem Cankaya and Anamaria Popescu provided essential help with data collection and entry; thank you. I am indebted to the Natural Sciences and Engineering Research Council and the Social Sciences and Humanities Research Council of Canada for providing the financial support necessary to conduct this research.

There are so many schools from the Black Gold and St. Albert Protestant school districts that opened their doors to me. Without you these ideas would not have been captured, the stories would not have been told, and I think we would have been poorer for it. Thank you for sharing your tight spaces, cramped schedules, and burdened workloads.

Finally, I would like to acknowledge my husband (Jon), my family, and my friends for all their support and encouragement over the years. You encouraged me to continue even when doing so meant I could not be as involved in your lives as I should have been. Thank you, and I love you.

Table of Contents

CHAPTER I: INTRODUCTION.. 1

- Limited Views of the Equal Sign.. 2
- The Relation between (Mis)Understanding the Equal Sign and Performance on Equivalence Problems.. 3
- Overcoming Difficulties Related to the Equal Sign and Equivalence........................ 4
- Effects of Arithmetic Contexts.. 9
- Children's Strategies and Emerging Competence: Studying Change........................ 13
- Teachers' Awareness of Equivalence Difficulties and Their Ideas for Improving Instruction.. 17
- Summary of the Proposed Research Studies.. 21
- References.. 23

CHAPTER II: LEARNING COMPLEX ARITHMETIC WITHOUT SYMBOLS: EXPLORING CHILDREN'S MISUNDERSTANDING OF THE EQUAL SIGN WITH NONSYMBOLIC EQUIVALENCE PROBLEMS 28

- Introduction.. 28
  - Operator and Relational Views of the Equal Sign.. 29
  - Evidence of Early Equivalence Skills.. 31
  - Mapping Understanding of Equivalence to the Equal Sign........................ 34

- Study 1.. 35
- Method.. 37
  - Participants and Design.. 37
  - Materials and Procedures.. 38
    - Practice problems.. 38
    - Test problems.. 39
    - Arithmetic problems.. 40
- Results.. 40
  - Overall Performance.. 40
  - Individual Differences.. 42
    - High-accuracy group.. 44
    - Zero-accuracy group.. 47
    - Low-accuracy group.. 48
- Discussion.. 51

- Study 2.. 53
- Method.. 54
  - Participants and Design.. 54
  - Materials and Procedures.. 55
    - Nonsymbolic condition.. 55
    - Symbolic condition.. 56
    - Arithmetic problems.. 57

      Verbal interpretation of the equal sign.. 58
  Results 58
    Strategy Use.. 61
    Verbal Interpretations of the Equal Sign.................................... 65
    Age-Related Changes in Performance.. 65
  Discussion... 67

  General Discussion.. 70
  References... 74

CHAPTER III: JUST WHERE DO THEY GO WRONG? EXPLORING YOUNG CHILDREN'S MISUNDERSTANDING OF TH EQUAL SIGN BY MANIPULATING THE DEGREE OF SYMBOLIZATION IN EQUIVALENCE PROBLEMS................. 79

  Introduction... 79
    Development of Symbol Use... 79
    Understanding the Symbol of Equivalence................................... 81
    Schemas of Typical Arithmetic Patterns....................................... 85
  Method... 88
    Participants and Design... 88
    Materials and Procedures.. 89
      Symbolic condition.. 89
      Semi-symbolic condition... 90
      Arithmetic problems.. 91
      Verbal interpretation of the equal sign...................................... 91
  Results and Discussion.. 92
    Accuracy... 92
    Strategy Use.. 95
    Verbal Interpretations of the Equal Sign....................................... 103
  General Discussion.. 106
  References... 110

CHAPTER IV: TRACKING THE PATHS OF CHANGE FROM ADHERENCE TO OPERATIONAL PATTERNS TO A RELATIONAL UNDERSTANDING OF THE EQUAL SIGN.. 113

  Introduction... 113
    Boundary Conditions of Conceptual Change................................ 113
    Operational Patterns and Success on Equivalence Problems........ 117
    Effects of Prior Experience.. 119
    Dimensions of Change... 121
    Maintenance of Learning and Flexibility of Thought.................... 122
  Method... 123
    Participants and Design... 123
    Materials and Procedures.. 124
      Screening test... 125

Operational patterns.. 128
Equivalence sessions.. 131
Clinical interview.. 136
Follow-up test.. 136
Classroom visit.. 136
Results and Discussion.. 138
Adherence to Addition Schemas.. 138
Initial performance.. 139
Predicting change on equivalence sessions.. 140
Changes in adherence to the addition schema.. 141
Effects of Prior Experience: Relations Between Exposure Condition and Performance on Equivalence Problems.. 146
Exposure problems.. 147
Symbolic test problems.. 147
Maintenance and transfer of equivalence skills.. 155
Five Dimensions of Change.. 161
Path.. 161
Rate.. 174
Breadth.. 176
Variability.. 178
Sources 178
Individual Differences.. 181
Clinical Interview.. 184
General Discussion.. 187
Adherence to Addition Schemas.. 188
Effects of Prior Experience.. 191
Maintaining and Transferring Knowledge.. 195
Five Dimensions of Change.. 197
Representing Mathematical Relations: Thinking Flexibly About Symbols.. 198
Summary.. 200
References.. 202

CHAPTER V: TEACHERS' PERCEPTIONS ABOUT CHILDRENS' UNDERSTANDING OF THE EQUAL SIGN.. 205

Introduction.. 205
Link between Teachers' Perceptions and Students' Performance.. 206
Method.. 209
Results and Discussion.. 210
Teachers' Perceptions of Children's View of the Equal Sign.. 210
Instructional Approaches and Teachers' Suggestions.. 213
Teachers' Estimates about Success on Equivalence Problems.. 218
General Discussion.. 223
References.. 231

CHAPTER VI: GENERAL DISCUSSION AND CONCLUSIONS.......................... 233

References.. 243

APPENDICES.. 244

List of Tables

4-1 The number of participating Grade 2 students in each session, and the mean number of intervening days between each session and the next... 126

4-2 Number of Grade 2 students in each condition who met the stop criterion and the last session completed.................................... 149

4-3 The number of Grade 2 children who received the feedback, and the session number on which they began receiving feedback for each of the three conditions.. 150

4-4 Number of children from each of the three conditions whose paths of change could be best described by one of five specific patterns........... 163

4-5 Correlations for the Grade 2 children who completed all equivalence testing sessions in the microgenetic study.................................... 183

List of Figures

2-1 Kindergarten children's accuracy on 10 arithmetic equivalence problems presented nonsymbolically.. 43

2-2 Distribution of matches between solution choices and justifications across the ten equivalence problems for all 31 kindergarten children. Matches were scored as one point, and mismatches zero, for a maximum agreement score of 10.. 46

2-3 Mean accuracy by Grade 2 children (symbolic/nonsymbolic and nonsymbolic/symbolic session order groups) on equivalence problems presented in both symbolic and nonsymbolic conditions, with standard error bars.. 59

2-4 Number of the 16 Grade 2 children in the symbolic/nonsymbolic group who used specific strategies on both symbolic and nonsymbolic sessions.. 63

2-5 Number of the 16 Grade 2 children in the nonsymbolic/symbolic group who used specific strategies on both symbolic and nonsymbolic sessions.. 64

3-1 Mean accuracy by Grade 2 children (symbolic/semi-symbolic and semi-symbolic/symbolic session order groups) on part-whole and combination equivalence problems presented in both symbolic and semi-symbolic conditions, with standard error bars... 94

3-2 Distribution of total agreement scores (with a maximum score of 32 matches between solutions and justifications) for the 32 Grade 2 children in the symbolic/semi-symbolic and semi-symbolic/symbolic conditions... 98

3-3 Percentage of strategy types, as defined by matches between solutions and justifications, on symbolic (top panel) and semi-symbolic (bottom panel) sessions by children in the symbolic/semi-symbolic group......... 100

3-4 Percentage of strategy types, as defined by matches between solutions and justifications, on symbolic (top panel) and semi-symbolic (bottom panel) sessions by children in the semi-symbolic/symbolic group......... 101

3-5 Percentage of strategy types on symbolic problems, as defined by matches between solutions and justifications, for children in the symbolic/nonsymbolic (top panel, Chapter II, Study 2) and symbolic/semi-symbolic (bottom panel) groups.............................. 104

3-6 Number of children from both the symbolic/nonsymbolic (Chapter II, Study 2) and symbolic/semi-symbolic groups who used specific strategies consistently, defined as choosing the same type of solution option for at least 13 of 16 problems, on the symbolic sessions............ 105

4-1 Average total scores (in percent) on the two operational patterns tests for Grade 2 children as a function of exposure conditions (with standard error bars).. 144

4-2 Grade 2 children's scores (in percent) on three of the operational patterns tasks from the first to second administration (with standard error bars).... 145

4-3 Mean number of strategies used per session on exposure and test phases of the equivalence testing sessions for all 34 Grade 2 children (with standard error bars).. 157

4-4 Percent correct on the maintenance and far transfer problems of the two-week follow-up test by all children (*N* = 32) who either met or did not meet the stop criterion in the three conditions (with standard error bars).. 158

4-5 Performance on the two-week follow-up test by the 20 children who met the stop criterion in either three to five sessions or six to nine sessions, regardless of exposure condition (with standard error bars).................. 160

4-6 "Joshua's" paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as for his strategy use (number of different strategies) and confidence ratings (maximum = 7) across the sessions. Joshua was in the nonsymbolic condition, and received feedback beginning on the fourth session............................. 166

4-7 "Tanya's" paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as for her strategy use (number of different strategies) and confidence ratings (maximum = 7) across the sessions. Tanya was in the semi-symbolic condition, and began to receive feedback in the sixth session................................ 169

4-8 "Olivia's" paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as her strategy confidence ratings (maximum = 7) across the sessions. Olivia was in the semi-symbolic condition, and began to receive feedback in the fourth session. Her paths of change represent low accuracy throughout the equivalence testing sessions and a decrease in confidence ratings........................ 171

4-9 “Peter’s” paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as his confidence ratings (maximum = 7) across the sessions. Peter was in the nonsymbolic condition, and began to receive feedback on the fifth and final session. He declined participation for any remaining sessions........................ 172

5-1 Categories of responses to first question: Would you say your students have a fairly good understanding of what the equal sign means? (*N* = 17).. 212

5-2 Categories of responses to the third question: Suppose you had to teach or remind your students about what the equal sign means. How would you go about doing that? (*N* = 17)... 216

5-3 Categories of responses to the fourth question: Have you used the equal sign in different contexts with your student (such as showing equal amounts of money, e.g., 5 pennies = 1 nickel), or primarily in the context of arithmetic questions? (*N* = 17).. 217

5-4 Categories of responses to the eighth question: If you think that some students in your class would fail, why do you think they would fail? (*N* = 17).. 222

CHAPTER 1

INTRODUCTION

A recent trend in the mathematical cognition and cognitive development literature is examination of the sophisticated strategies and concepts of very young children on various mathematical tasks. Numerous lines of research now point to levels of skill use and strategy variability previously attributed to much older children. Such research includes examination of the early emergence of mathematical concepts (Rasmussen, Ho, & Bisanz, 2003; Sherman & Bisanz, in press), counting and arithmetic skills (Baroody, 1987; Bisanz, Sherman, Rasmussen, & Ho, 2005; Siegler & Jenkins, 1989; Wynn, 1992, 2000), variability and strategy choices (Shrager & Siegler, 1998; Siegler & Shipley, 1995; Siegler & Shrager, 1984), and numeric versus non-numerical internal representations (Kahneman, Treisman, & Gibbs, 1992; Mix, Huttenlocher, & Levine, 2002; Simon, 1997). Children's emerging understanding of underlying principles of arithmetic, such as commutativity (Baroody, Wilkins, & Tiilikainen, 2003) and inversion, has been examined using various approaches. Using a nonverbal format, for example, Sherman and Bisanz (in press) were able to assess three-year-olds' sensitivity to the principle of inversion (that $a + b - b$ *must* equal a). Surprisingly, the preschoolers demonstrated a convincing advantage for inversion problems compared to similar standard problems (e.g., $a + b - c$), long before proficiency in counting or instruction in formal arithmetic.

Despite some research suggesting remarkable competency at young ages, there is at least one line of research examining an opposite phenomenon, specifically, elementary school-aged children's overwhelming failure on arithmetic equivalence problems (e.g., 4

+ 2 + 6 = 3 + __). In fact, almost 90% of fourth- and fifth-grade students fail such problems (Perry, 1985, as cited in Perry, Church & Goldin-Meadow, 1988; for a notable exception, see Chapter IV below), and failure is typically attributed to a misunderstanding of the equal sign (Alibali, 1999; Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; McNeil & Alibali, 2005a; McNeil & Alibali, 2004; Perry et al., 1988). Success on arithmetic equivalence problems relies on understanding that both sides of the equal sign must be quantitatively identical. Although elementary school-aged children have little difficulty with the arithmetic involved in solving problems of this magnitude, they overwhelmingly fail to use arithmetic to make both sides of the equal sign equivalent (Alibali, 1999; McNeil & Alibali, 2005a, 2005b; McNeil & Alibali, 2004; Perry et al., 1988). Understanding the equal sign in a fully *relational* way (i.e., both sides of the equal sign must be equivalent) rather than in an *operator* way (i.e., the equal sign means that operations are needed to solve, such as adding up all terms in 4 + 2 + 6 = 3 + __ and putting 15 on the line) is believed to be essential for higher mathematics such as algebra (Kieran, 1981; Knuth et al., 2005; McNeil & Alibali, 2005a, 2005b).

Limited Views of the Equal Sign

Several researchers have suggested that children have a limited, operator interpretation of the equal sign, particularly in arithmetic contexts. Seo and Ginsburg (2003) presented Grade 2 children with various expressions in which the equal sign was present, including coins and Cuisenaire rods (e.g., 1 dime = 10 pennies, or one red rod = 2 white rods), canonical (e.g., 2 + 3 = __) and noncanonical (e.g., __ = 2 + 3) arithmetic expressions, and no context (e.g., asking children what the equal sign means). For all expressions, children were asked to explain what the equal sign means. Children were more likely to give a relational response when the equal sign was in the context of coins or Cuisenaire rods than when it was in the context of arithmetic. Seo and Ginsburg concluded that children hold both interpretations of the equal sign, yet how they view the symbol on any given problem is influenced by the context in which they find the symbol.

Knuth et al. (2005) also found that children have a poor understanding of the equal sign in arithmetic contexts. Three hundred and seventy-three students from Grades 6 to 8 were given the expression “3 + 4 = 7” and asked to explain what “=”means. Even among the Grade 8 children, only 46% gave a relational response. Sherman and Bisanz (2004) also found an overwhelming tendency among young children to give operator responses when asked what the equal sign meant in symbolic arithmetic contexts. In fact, 94% of Grade 2 children’s responses were consistent with an operator interpretation. Such a limited understanding of the equal sign appears to negatively impact performance on arithmetic equivalence problems, and has also been implicated as a major contributor to difficulties in higher mathematics (Kieran, 1981; Knuth et al., 2005; Knuth, Stephens, McNeil, & Alibali, 2006; McNeil & Alibali, 2005a, 2005b).

The Relation Between (Mis)Understanding the Equal Sign and Performance on Equivalence Problems

Is children’s failure on arithmetic equivalence problems due to their operator view of the equal sign? It appears that way, as Knuth et al. (2005) found a relation between Grade 6 to 8 students’ definition of the equal sign and their performance on a test of equivalence using equivalence problems. Specifically, students saw “3 + 4 = 7” and were asked what the name of the symbol was, what it meant, and if the symbol could mean anything else. Responses were coded as being relational, operational, or other. They were then shown “2x ☐ + 15 = 31” and “2x ☐ + 15 – 9 = 31 - 9” and asked if the number that goes in the box was the same number in both equations and to explain their answers. Responses were coded as answer-after-equal-sign, recognize equivalence, solve and compare, or other. Students who provided a relational interpretation of the equal sign were more likely to realize that the two equations had the same solutions without having to solve and compare than were children who gave an operator interpretation. Thus, how children view the equal sign appears to directly influence how they approach problems where there are terms on both sides of the equal sign. The link between children’s misunderstanding of the equal sign and compromised performance on certain types of equations holds even when general mathematics ability is controlled (Knuth et al., 2006). An operator interpretation of the equal sign seems to limit children’s ability to think about equivalence problems as such.

Overcoming Difficulties Related to the Equal Sign and Equivalence

How can we help children surmount this overwhelming difficulty with equivalence problems and misunderstanding the equal sign? Several intervention or teaching studies have been conducted with the goal of finding the best ways to improve children’s performance on arithmetic equivalence problems (Alibali, 1999; McNeil & Alibali, 2000; Perry, 1991). Unfortunately, most techniques have been met with limited success, and long-term effects have not been assessed. For example, most Grade 3 and 4 students were still unsuccessful on equivalence problems even after considering the analogy of balancing a teeter-totter or direct instruction about the principle of equivalence (Alibali, 1999).

Despite convincing results showing that children consistently fail equivalence problems and give operational interpretations of the equal sign, typical tests of equivalence do not allow researchers to distinguish between children's misconceptions of the equal sign and their recognition (and use) of appropriate procedures. That is, all previous tests of equivalence have been given using mathematical symbols, making it difficult to assess whether children's failure is due to (a) the presence of symbols (such as the equal sign) that may cause incorrect strategies or (b) difficulties using arithmetic to make both sides of the equal sign quantitatively identical. To explore the possible differences between these two factors (symbols versus arithmetic), a novel context was devised to assess children's performance on equivalence problems without any mathematical symbols (Sherman & Bisanz, 2004). Performance on equivalence problems in this *nonsymbolic* format indicated that Grade 2 children, and even some kindergarten children, are indeed capable of using arithmetic to make sums on both sides of the equal sign equivalent. This was the first definitive evidence that children's poor performance on equivalence problems is limited to symbolic contexts, as well as the fact that children may enter school with the ability to solve equivalence problems. Unfortunately, children's early knowledge of equivalence, even using arithmetic to ensure equivalence, may not be effectively mapped onto typical symbolic problems taught in school. Specifically, children may not be linking their ability to reason about arithmetic equivalence to the symbol of equivalence (=) as taught in standard mathematical curricula.

Several studies have shown that adapted curricula, or programs specifically designed to teach students about the equal sign in a relational manner, can be effective, although such programs are extremely rare. In a year-long socio-constructivist examination of children's understanding of equivalence, Sáenz-Ludlow and Walgamuth (1998) documented the extensive amount of discussion and time it took for class of Grade 3 students to finally view the equal sign as a relational symbol in arithmetic contexts. Fourteen students were videotaped daily, and samples of their math notes and solutions were analyzed, to document the slow change from a primarily operator view of the equal sign to a relational understanding by the end of the school year. After numerous class discussions about problems such as 246 + 14 = __ + 246 and the role of the equal sign, children finally made the connection when invited to think about "=" and equivalence as being "even Steven," a colloquialism that allowed students to realize the relational nature of the equal sign. Baroody and Ginsburg (1983) examined children's view of the equal sign in a school with the Wynroth program, a systematic approach to teaching the relational meaning of the equivalence sign. They too found that, with deliberate and prolonged instruction, even Grade 2 children demonstrated a "tentative grasp of a relational view" (p. 207). Unfortunately, such programs are extremely rare and do not reflect the current status of North American classrooms, as evident in the overwhelming failure on equivalence problems by students in the many studies done (Alibali, 1999; Charchun, Bisanz, & Sherman, 2006; Knuth et al., 2005; McNeil & Alibali, 2005a; McNeil & Alibali, 2004; Perry, 1985, as cited in Perry et al., 1988; Sherman & Bisanz, 2004; Watchorn & Bisanz, 2005).

Given the well-documented difficulty with equivalence problems and the extensive measures that have been undertaken to correct children's misunderstanding of the equal sign, the impressive accuracy levels among young children on nonsymbolic problems is appealing from both psychological and educational standpoints. Grade 2 children's overwhelming success on nonsymbolic problems (Sherman & Bisanz, 2004) raises the question of whether practice with equivalence problems in this condition might be useful for successfully solving symbolic equivalence problems. That is, children may map their strategies for making both sides of the nonsymbolic equations equal to symbolic problems. Unfortunately, these results came from a study with a between-subject design, and therefore the effects of exposure to nonsymbolic problems on the accuracy of symbolic equivalence problems could not be assessed. Therefore, an important question remained: Could experience with nonsymbolic equivalence problems help children correctly solve symbolic equivalence problems? In other words, could exposure to problems without the equal sign (or other symbols that placed the problems in an obvious arithmetic context) help children succeed on the same problems presented with symbols? To examine these questions, I conducted a within-subject design with children in Grade 2 to replicate and extend the previous study in examining potential order effects of solving equivalence problems with and without symbols (Chapter II). Furthermore, in the previous study (Sherman & Bisanz, 2004), children responded to equivalence problems by producing solutions from an unrestricted range. Thus, accuracy on equivalence problems could not be compared to chance levels, which would have been useful for determining whether low accuracy by both the kindergarten children and Grade 2 children in the symbolic group nevertheless provided some evidence for understanding equivalence.

In light of the previous study's limitations, I tested both kindergarten children (Chapter II, Study 1) and Grade 2 children (Chapter II, Study 2) using a multiple-choice paradigm so that accuracy can be compared to chance levels. Two main questions are addressed in these studies:(a) Can exposure to nonsymbolic problems help performance on subsequent symbolic equivalence problems, and (b) Do children begin formal mathematical instruction already having the ability to make two sides of an equation equivalent? Addressing both these questions is important for understanding the process of mapping. In the first study I explored mapping by establishing whether children enter school already having the ability to use arithmetic to make both sides of an equation equivalent. If kindergarten children demonstrate understanding of arithmetic equivalence problems, and enter school with this sophisticated skill set, then children may be experiencing difficulty mapping formally instructed arithmetic concepts and goals in symbolic contexts to their understanding of equivalence in nonsymbolic contexts.

In the second study I examined whether children benefit from exposure to nonsymbolic problems *before* exposure to symbolic problems, such that children's success on nonsymbolic problems may help them map their strategies onto subsequent symbolic problems. In this case, making explicit links between nonsymbolically and symbolically presented problems may help children to improve performance on equivalence problems *and* to have a more relational view of the equal sign. The multiple-choice method is better suited in both studies for determining whether children's understanding of arithmetic equivalence develops *prior* to formal instruction than in previous research with unrestricted ranges of responses. That is, with the multiple-choice method accuracy levels can be compared to chance levels to determine whether even very young children show sensitivity to the demands of these complex arithmetic equivalence problems.

Effects of Arithmetic Contexts

The studies in Chapter II with the nonsymbolic format are important and useful for examining children's ability to solve equivalence problems apart from the symbols (i.e., the equal sign) that may lead to incorrect strategies, such as adding all terms together. Why would symbols (or perhaps just one symbol in particular) cause children to use incorrect strategies when solving problems? McNeil and Alibali (2004, 2005b) suggested a *change resistance account* for explaining the difficulty with symbolic equivalence problems. This account, which focuses on what children already know or do as opposed to what they *lack*, posits that children's failure on equivalence problems arises, at least somewhat, from "their early and elongated experience with arithmetic operations" (McNeil & Alibali, 2005b, p. 4). Experience with symbolic canonical arithmetic problems ($a + b =$ c) may lead children to form patterns about operating from left to right to solve, and this adherence to operational patterns may result in failure on problems that differ from canonical forms (such as equivalence problems, $a + b = c +$ __). The operational patterns include: (a) a *strategy* of operating on all given numbers (i.e., add up all terms, regardless of where they are in the expression), (b) a *perceptual pattern* whereby solutions always immediately follow the equal sign, and (c) an incorrect *concept* of the equal sign as meaning "the answer comes next." These operational patterns form an *addition schema*, which in turn influences how children attend, encode, and interpret novel arithmetic problems. Children with strong addition schemas may encode and interpret an equivalence problem, such as "3 + 2 + 4 = 1 + __," as "3 + 2 + 4 + 1 = __" so that the problem fits with their belief that the answer must follow the equal sign. The change resistance account makes the counterintuitive hypothesis that children with more experience with canonical arithmetic problems, and therefore stronger addition schemas, will have lower accuracy on equivalence problems.

McNeil and Alibali (2005b) tested the change resistance account in two studies. In the first they examined individual differences in 7- to 11-years-olds on tests of operational patterns and a test of learning to solve equivalence problems. They found that the more children adhered to operational patterns, the less likely they were to learn from brief lessons for solving equivalence problems. In the second study, they explored the effects of activating operational patterns, rather than using correlational data, to examine the link between adherence to operational patterns, or an addition schema, and difficulties with equivalence problems. Undergraduate students were randomly assigned to receiving either stimuli designed to activate the addition schema or neutral stimuli, followed by an equivalence problem-solving phase. In the activation phase, students were asked to press a "yes" or "no" key to indicate whether a set of five stimuli matched the target. There were 24 trials in total, including eight trails designed to activate each of the three aspects of operational patterns: the strategy of adding all numbers together, the perceptual pattern of answers following the equal sign, and the concept of the equal sign meaning the total. In the neutral phase, students saw words that were not related to arithmetic. In the problem-solving phase, students were asked to solve two canonical arithmetic problems, followed by eight equivalence problems.

Almost all (92%) undergraduate students in the neutral control group solved at least seven of the eight equivalence problems correctly, but the corresponding rate in the activation group was an astonishingly low 42%. Data from both studies support the change resistance account, as students who either adhered to operational patterns more

(McNeil & Alibali, 2005b, Study 1) or who had operational patterns activated (McNeil & Alibali, 2005b, Study 2) appeared to suffer when faced with equivalence tasks. Perhaps more importantly, experience with typical canonical arithmetic problems, specifically those presented with mathematical symbols, can hinder performance on complex arithmetic problems, rather than help performance. The authors point out that this counterintuitive hypothesis raises caution about the tenet that practice makes perfect.

One important component of operational patterns is the prevalent (and incorrect) concept that the equal sign indicates the sum or solution to the problem. Knuth et al. (2005) found a direct relation between interpretations of the equal sign in particular and performance with equivalence problems. Unfortunately, almost all previous research examining children's misunderstanding of the equal sign and failure on equivalence problems uses entirely symbolic formats (except Sherman & Bisanz, 2004), and inferences from these symbolic problems are made about children's misunderstanding of one symbol in isolation: the equal sign. It is possible, however, that a combination of specific mathematical symbols (or all of them together) leads to poor performance on equivalence problems, rather than just the equal sign per se. Because previous research has not addressed this question, with the exception of a completely nonsymbolic format (Sherman & Bisanz, 2004), I examined children's performance on equivalence problems that only contain some aspects of typical symbolic problems (Chapter III). In a novel *semi-symbolic* format, I investigated children's ability to make two sides of an equation quantitatively identical in a context in which some symbols ("+" and "=") exist, yet other mathematical symbols are substituted (e.g., "••" for "2"). Admittedly, additional combinations of conventional and unconventional mathematical symbols would be essential to pinpoint the precise degree of symbolization (if one exists) necessary to activate children's addition schema. However, this study is an important first step toward examining children's sensitivity to a context that may or may not look sufficiently like symbolic arithmetic.

We know from Sherman and Bisanz (2004) that in the absence of all mathematical symbols Grade 2 children and even some kindergarten children have high accuracy, yet in the presence of all symbols performance is extremely low. This new format helps to narrow the number or combination of mathematical symbols that is necessary to induce an addition schema. If children's accuracy is low on semi-symbolic problems, it may be because even though these problems don't look like typical arithmetic problems, the equal sign is still present and leads to activation of the disruptive addition schema. In contrast, if children's accuracy on these problems is high, then perhaps the equal sign alone is not enough to cause the difficulty, but rather the broader arithmetic context when all parts of the equation are represented using symbols. As in Study 2 of Chapter II, exposure to the less symbolic problems first (in this case, semi-symbolic problems) may positively influence accuracy on subsequent symbolic equivalence problems. I examined the questions above using a within-subject design and multiple-choice responses, as in the first study. Thus, in Chapter III, I explored the change resistance account of operational patterns by altering the number of symbols children typically see using a new semi-symbolic format.

Children's Strategies and Emerging Competence: Studying Change

If the hypotheses from studies in Chapters II and III are correct, then it would appear that children's accuracy on arithmetic equivalence problems could be highly modifiable depending on the order in which they receive problems in the various contexts (i.e., symbolic, nonsymbolic, and semi-symbolic). In this case, children's view of the equal sign, or at least their performance on equivalence problems, would be amenable to change. I would not be able to discern, however, much about the nature of change, such as how quickly change occurs or the particular experiences that precede change, based on observing children across two sessions. Siegler (1996) described five unique dimensions of change that are observable using microgenetic methods, including the path, rate, breadth, variability, and sources of change. Change on equivalence problems can be operationalized as changes in (a) children's accuracy and various strategies for solving equivalence problems, such as add-all and relational approaches, (b) their justifications for their strategies and solutions, and (c) the confidence they have in their solutions and strategies.

Using these diverse measures of change in a microgenetic design allows me to examine all five aspects of change. The path of change, which can be operationalized as the sequences of qualitatively distinct understandings, can be examined by tracking the various ways in which children solve equivalence problems and define the equal sign. Because I wish to examine change in performance and conceptual understanding over time, I tested children who initially demonstrate high accuracy on canonical problems but poor performance on equivalence problems. Children in Grades 2 and 4 received a screening test to find participants who meet the performance criteria. Then I randomly assigned students to one of three conditions--nonsymbolic, semi-symbolic, and symbolic only--and tracked their accuracy on symbolic equivalence problems as well as their views of the equal sign and verbal reports until they consistently solved problems relationally. I expected that children, at least those in the nonsymbolic and semi-symbolic groups, would progress from one or more operator strategies, such as add-all, to a relational strategy (i.e., using arithmetic to make both sides of the equal sign equivalent). Examining the path is important for determining whether predictable sequences of strategy change exist, how long distinct strategies typically exist, and whether important deviations to sequence occur (and if so, why). Details about paths of change are necessary before modifications to instruction can be proposed. For example, using semi-symbolic contexts to teach children about equivalence may be beneficial for later symbolic problems in terms of high accuracy levels, yet the path of change using semi-symbolic problems first may involve more intermediate steps susceptible to delays or deviations than practice with some other intervention-type context such as nonsymbolic problems.

Examining the rate of change, or the speed at which changes in strategy use or interpretations occur, can also include assessing potential experiences that affect the rate of changes. Some children may adhere to particular strategies and definitions longer than other children, and differences between the rate of change among children may relate to their grade, experimental group (nonsymbolic, semi-symbolic, or symbolic), or their adherence to the addition schema. Determining factors that affect the rate of change is vital for optimizing instruction or predicting success on equivalence problems. For example, if exposure to nonsymbolic problems increases the rate of change from low to high accuracy levels on symbolic problems relative to the other two conditions, then teachers may wish to employ nonsymbolic problems in the classroom to optimize or speed the development of relational understanding.

Breadth of change in this study would include assessing whether children's changes in strategy use extends to equivalence problems types that have terms arranged in various positions (e.g., combination problems such as 3 + 4 + 2 = 3 + __ versus part-whole problems such as 5 + 4 = 2 + __; see Sherman & Bisanz, 2004) and are presented in different contexts (e.g., both nonsymbolic *and* symbolic). For example, children may progress from operator to a relational strategy in the nonsymbolic context yet continue to solve problems operationally in the symbolic context. The breadth of change, in this example, would not be as broad as in children who solve problems relationally in both contexts and across different types. Strategy change may be staggered rather than all-or-none, such that children change in some conditions before they change in other conditions. Children's ability to generalize their change across formats and contexts is important for understanding how children map their understanding across situations. Differences in breadth of change can be informative for instructional practices, such that some children may need deliberate instruction linking strategies and interpretations across contexts and problem types.

The goals of examining the path, rate, and breadth of change imply the existence of variability such that children may differ from each other in terms of the strategies and interpretations expressed, the rate at which strategies change, and factors that influence the generalization of change across different contexts and problem types. However, as Siegler and Shipley (1995) pointed out, variability in strategy use often exists within children, not just across children, such that the same child may demonstrate multiple strategies for solving the same type of problem. The microgenetic design will be useful for examining variability across children, such as whether children in one group use different or more strategies than children in another group, as well as within children. Individual trajectories can be established to compare the variability within each child and to explore potential patterns of variability that may relate to group status, grade, or adherence to the addition schema.

Assuming that children will change their strategy use and interpretations of the equal sign begs the question of *why* changes occur, or what the sources of change are. For example, does change occur *after* experience with nonsymbolic problems but *not* after experience with symbolic problems? Identifying sources of changes has direct instructional implications. If one of the contexts (i.e., nonsymbolic) is convincingly associated with high accuracy on symbolic equivalence problems, then teachers can implement experience with problems in that context in their classrooms. Identifying experiences that directly precede high levels of accuracy on symbolic equivalence problems can help isolate potential sources of change in children's strategy use or view of the equal sign.

In addition to examining the five dimensions of change, microgenetic studies are useful for identifying experiences that might "contribute to cognitive change [by] providing intense exposure to such experiences and observing children's reactions to them" (Siegler, 1995, p. 233). In Chapter IV, I examined how experience with nonsymbolic and semi-symbolic equivalence problems affects children's performance on symbolic problems in more detail than was possible in the studies in Chapters II and III.

In this way I was able to learn more about how children's experience in the specific contexts change their accuracy, judgments and explanations, as well as identify potential methods for teaching elementary school children about the equal sign and equivalence.

Teachers' Awareness of Equivalence Difficulties and Their Ideas for Improving Instruction

Despite the overwhelming failure on arithmetic equivalence problems, and children's persistent misunderstanding of the equal sign, there have been few studies examining teachers' awareness of the problem. Research has shown a link between teachers' knowledge of student understanding and students' actual performance in elementary topics such as whole numbers (Carpenter, Fennema, Peterson, Chiage, & Loef, 1989), and highlighted the important relation between teacher knowledge and student achievement. "Because the beliefs that teachers hold are so instrumental in shaping mathematics teachers' decisions and actions, it is important that these beliefs be a focus of educational research; in addition, because these decisions affect students' learning experiences so directly, it is equally important to understand the accuracy of teachers' beliefs" (Nathan & Koedinger, 2000a, p. 210).

Some researchers have noted that when teachers are unaware of the source of students' errors or confusion, they may focus instruction in ways that do not meet students' actual needs (Nathan & Petrosino, 2003). An "expert blind spot" (EBS), which arises when experts in a specific area or field are unaware of the learning and developmental needs of non-experts, may apply to school teachers in areas such as arithmetic and algebra (Nathan & Petrosino, 2003). Teachers with advanced knowledge of arithmetic, for example, may misattribute children's difficulty with equivalence problems to a deficit in equation mastery, rather than a misunderstanding of the equal sign. Excessive practice on canonical arithmetic, such as "drill and kill" exercises (English & Halford, 1995), with the hopes of improving equation mastery, may indeed have the reverse effect, as predicted by the change resistance account.

Nathan and Koedinger (2000a) proposed the *symbol-precedence* hypothesis to explain why many high school teachers overestimate students' competence on symbolic algebra problems and underestimate their performance on word- or story- problem algebra problems (which the authors referred to as verbal formats). That is, because mathematics textbooks tend to focus first on symbols and symbol manipulation and only later introduce verbal problems as applications of symbolic problems, teachers often believe that students are highly competent with mathematical symbols and will have higher accuracy on symbolic than verbal algebra and arithmetic problems. For example, Nathan and Koedinger (2000b) asked 67 mathematics teachers (Grades 7-12) to rank order (from easiest to hardest) 12 problems. The problems were created by combining two types (canonical arithmetic or *result unknown*, and algebraic or *start unknown*) and three presentation formats (symbolic, word, and story). Most of the teachers ranked algebra problems as more difficult than canonical arithmetic problems, regardless of presentation format. Furthermore, high school teachers often agreed with the symbol-precedence view, rating symbolic equations as easier than the other problems, including verbal problems. Students' actual data confirmed that presentation format had a significant impact on performance, but not in the order predicted by many of the teachers. Instead, students' performance demonstrated that symbolically presented problems were the most difficult. Despite teachers' prediction about students' proficiency with symbols and weakness with verbal problems, children may have highly conceptual and informal strategies to solve arithmetic and algebra problems before they have an accurate understanding of the actual symbols in algebra (such as the equal sign and variables).

Asquith, Stephens, Knuth, and Alibali (under review) interviewed twenty middle-school (Grades 6-8) teachers to determine their awareness of students' understanding of two algebraic concepts, the equal sign and variable. Student data were collected from Grades 6-8 at a different school within the same district as the teachers. The student worksheet was the same as in the study conducted by Knuth et al. (2005) and described above. Teachers were shown the student worksheet (in which students were asked to define the equal sign, determine whether two equations were equivalent, define what a variable is, and decide which algebraic expression was larger), and asked three questions: (a) what answers (correct or not) would you expect your students to give and what strategies might they have used, (b) how many students from across the district would get this problem correct and how many would use each strategy, and (c) could you explain your reasoning for your answers? Although students tend to progress from an operator to a relational view of the equal sign, they do not do so as quickly as their teachers predict. For example, teachers predicted that 73% of the Grade 7 students would give a relational interpretation of the equal sign, even though only 37% actually did so. The researchers also noted that teachers may not be aware of the relation between students' understanding of the equal sign and their performance on algebraic problems such as judging the equivalence of different equations.

Anecdotal evidence from conversations I have had with elementary school teachers in the past suggest that teachers are not aware of children's misconceptions of the equal sign and believe that their students have a fully relational understanding. One teacher described the various contexts in which she instructed children about the equal sign, including language lessons, such as "little = small." Ironically, children's correct interpretations of the equal sign in such situations should not be taken as evidence of a relational, generalized, understanding of the equal sign. In fact, children often give relational interpretations in such contexts, yet continue to view the equal sign as an operator symbol in arithmetic contexts (McNeil & Alibali, 2005a; Seo & Ginsburg, 2003). Because teachers may be unaware that misinterpretations of the equal sign and equivalence problems exist, and because a correct, relational interpretation of the equal sign is necessary for success in higher mathematics such as algebra (Knuth et al., 2005; McNeil & Alibali, 2005a, 2005b), and finally because we need to know what teachers' believe about their students' understanding in order to promote change in the classroom, I examined teachers' views about their students' understanding of the equal sign and their instructional techniques in Chapter IV.

Rather than ask teachers to rank order various problem presentations and compare with actual performance, I presented teachers with a survey in which I asked much more directly, and in more detail, about teachers' beliefs about symbolic contexts, particularly about students' understanding of the equal sign and success on equivalence problems. If teachers' responses indicate that they are not aware of students' misunderstanding of the equal sign, then this would provide evidence that teachers are not receiving the message from the equivalence research and that perhaps a method for delivering this message should be established. If teachers do appear aware of children's misunderstanding of the equal sign, then further research will be necessary to establish a more systematic relation between awareness and students' performance on equivalence problems.

Summary of the Proposed Research Studies

In summary, in Chapter II I examined accuracy on nonsymbolic problems (kindergarten, Study 1, and Grade 2, Study 2) compared to symbolic problems (Grade 2 only), much as in Sherman and Bisanz (2004) except using a within-subject design that allowed me to determine whether there are order effects (Study 2). That is, exposure to nonsymbolic problems first may help children map the equivalence relation successfully onto subsequent symbolic problems. In both studies I used a multiple-choice design so that Grade 2 and kindergarten children's accuracy can be compared to chance levels. Previous research (Sherman & Bisanz, 2004) was not designed to definitively determine whether kindergarten children could solve nonsymbolic equivalence problems because of low accuracy on no comparison chance levels, so Study 2 was expected to extend previous results with more conclusive evidence regarding early equivalence ability. In Chapter III I studied Grade 2 children's performance on a novel, semi-symbolic format to more closely examine whether previously reported failure on equivalence problems also occurs when only some of the typical mathematical symbols are present (Study 3). I also used a within-subject design to determine whether an order effect occurs, to determine whether solving semi-symbolic problems first helps children map potential success onto subsequent symbolic problems.

In Chapter IV, I closely examined the nature of change in relation to exposure to either nonsymbolic or semi-symbolic problems on standard symbolic equivalence problems. A microgenetic design was used to monitor the potential changes in strategies, judgments, and confidence levels, perhaps as some Grade 2 and 4 children progress from being exclusively operational to, ideally, consistently relational. Finally, in Chapter IV I explored teachers' views about children's misunderstanding of the equal sign, how the equal sign is instructed in the classroom (or not), and what teachers propose would be ideal ways for instructing children about this symbol.

In each chapter I review the literature related to the specific questions addressed and link each chapter topic to the more general theme of this investigation. Children's conceptual views of the equal sign, their ability to map concepts from one context to another, their trajectories of change, and the views of their instructors are all pieces of some important developmental and educational questions, specifically, how do children learn to reason about symbols, what strategies do they use to solve complex and novel problems, and what type of supports or instruction can facilitate learning the best?

References

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127-145.

Asquith, P., Stephens, A. C., Knuth, E. J., & Alibali, M. W. (under review). Middle school mathematics teachers' knowledge of students' understanding of core algebraic concepts: Equal sign and variable. *Mathematics Teaching and Learning*.

Baroody, A. J. (1987). The development of counting strategies for single-digit addition. *Journal for Research in mathematics Education, 18*, 141-157.

Baroody, A. J., & Ginsburg, H.P. (1983). The effects of instruction on children's understanding of the "equals" sign. *The Elementary School Journal, 84*, 199-212.

Baroody, A. J., Wilkins, J. L., & Tiilikainen, S. H. (2003). The development of children's understanding of additive commutativity: From protoquantitative concept to general concept? In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 127-160). Mahwah, NJ: Erlbaum.

Bisanz, J., Sherman, J. L., Rasmussen, C., & Ho, E. (2005). Development of arithmetic skills and knowledge in preschool children. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition (pp. 143-162)*. New York: Psychology Press.

Carpenter, T.P., Fennema, E., Peterson, P.L., Chiage, P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational research journal, 2*6, 499-531.

Charchun, J., Bisanz, J., & Sherman, J. (2006, June). *Young children's interpretation of the equal sign: The obstinate operator bias*. Poster presented at the annual meeting of the Canadian Psychological Association, Calgary, AB.

English, L. D., & Halford, G. S. (1995). *Mathematics education: Models and processes*. Mahwah, NJ: Erlbaum and Associates.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175-219.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*, 317-326.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equality and variable. *Zentralblatt für Didaktik der Mathematik – International Reviews on Mathematical Education, 37*, 1-9.

Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education, 37*, 297-312.

McNeil, N. M., & Alibali, M. W. (2005a). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development, 6*, 285-306.

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology, 92*, 734-744.

Mix, K. S., Huttenlocher, J., & Levine, S., C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin, 128*, 278-294.

Nathan, M. J., & Koedinger, K. R. (2000a). An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction, 18*, 209-237.

Nathan, M. J., & Koedinger, K. R. (2000b). Teachers' and researchers' beliefs about the development of algebraic reasoning. *Journal for Research in Mathematics Education, 31*, 168- 191.

Nathan, M. J., & Petrosino, A. (2003). Expert blind spot among preservice teachers. *American Educational Research Journal, 40*, 905-928.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6*, 449-468.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Rasmussen, C., Ho, E., & Bisanz, J. (2004). Use of the mathematical principle of inversion in young children. *Journal of Experimental Child Psychology, 85*, 89-102.

Sáenz-Ludlow, A., & Walgamuth, C. (1998). Third graders' interpretations of equality and the equal symbol. *Educational Studies in Mathematics, 35*, 153-187.

Seo, K., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence…": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Sherman, J., & Bisanz, J. (in press). Evidence for mathematical inversion in three-year-olds. *Journal of Cognition and Development*.

Sherman, J., & Bisanz, J. (2004, May). *Symbolic interpretation: Examining children's view of the equal sign*. Poster presented at the 16th Annual American Psychological Society, Chicago, Illinois.

Shrager J., & Siegler, R.S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405-410.

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28*, 225-273.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Lawrence Erlbaum Associates.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skill* (pp. 229-294). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A "non-numerical" account. *Cognitive Development, 12*, 349-372.

Watchorn, R., & Bisanz, J. (2005, October). *Enduring instructional effects on equivalence problems*. Poster presented at the biennial meeting of the Cognitive Development Society, San Diego, CA.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

Wynn, K. (2000). Findings of addition and subtraction in infants are robust and consistent: Reply to Wakeley, Rivera, and Langer. *Child Development, 71,* 1535-1536.

CHAPTER II

LEARNING COMPLEX ARITHMETIC WITHOUT SYMBOLS: EXPLORING CHILDREN'S MISUNDERSTANDING OF THE EQUAL SIGN WITH NONSYMBOLIC EQUIVALENCE PROBLEMS

One theme that permeates research in mathematical development is how children's strategies and concepts progress in developmental sequence, how early abilities relate to later skills, and how deficits in certain domains impact performance on related areas. For example, early addition strategies, such as counting on fingers or using objects, is believed to relate to, and indeed establish fundamental connections for, later strategies such as retrieval (Bisanz, Sherman, Rasmussen, & Ho, 2004; Shrager & Siegler, 1998; Siegler & Shipley, 1995; Siegler & Shrager, 1984). Unfortunately, not much is known about children's early strategies for, and understanding of, arithmetic equivalence problems. Researchers have found that, in general, children in elementary school overwhelmingly fail tests of arithmetic equivalence (e.g., 4 + 2 + 6 = 3 + __) (Alibali, 1999; McNeil & Alibali, 2005a, 2005b; McNeil & Alibali, 2004; Perry, Church, & Goldin-Meadow, 1988; Sherman & Bisanz, 2004), yet children appear to enter school with an appreciation of equivalence in a general sense (i.e., children correctly match three puppet jumps to three rather than two objects) (Mix, 1999). Less clear is whether children enter school with an understanding of equivalence in more complex situations, such as arithmetic, and if so, how children map understanding of equivalence onto problems in the context of arithmetic, particularly problems presented with conventional symbols such as the equal sign.

Operator and Relational Views of the Equal Sign

When students in elementary school see a problem such as 4 + 2 + 6 = 3 + __, they overwhelmingly and erroneously solve the problem by adding up all terms and putting the total sum on the line (15), or by ignoring the last term (3) and putting the sum of the left side of the expression on the line (12). Although elementary school-aged children have little difficulty with the addition and subtraction involved in solving problems of this magnitude, they overwhelmingly fail to use arithmetic to make both sides of the equal sign equivalent (Alibali, 1999; McNeil & Alibali, 2005a; McNeil & Alibali, 2004; Perry et al., 1988; Sherman & Bisanz, 2004). Many researchers believe that it is children's misunderstanding of the equal sign that leads them to incorrect rather than mathematically sound strategies for solving equivalence problems (Baroody & Ginsburg, 1983; Kieran, 1981; McNeil & Alibali, 2004; Seo & Ginsburg, 2003). Success on such problems relies on understanding that both sides of the equal sign must be quantitatively identical.

Understanding the equal sign in a fully *relational* way (i.e., both sides of the equal sign must be equivalent) rather than in an *operator* way (i.e., the equal sign means that operations need to be executed in a conventional manner, such as adding all terms in the equation) is believed to be essential for higher mathematics such as algebra (Kieran, 1981; Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; McNeil & Alibali, 2005b). Young students often fail to view the equal sign in a relational manner, at least not when the equal sign occurs in the context of typical, symbolic (i.e., "5 + 2 = __") arithmetic problems. Seo and Ginsburg (2003) found that Grade 2 students gave operator interpretations of the equal sign, such as "the answer comes next", or "add up all the numbers", when the equal sign was presented in the context of arithmetic problems. In contrast, the students gave relational responses when the equal sign was presented in non-arithmetic contexts such as comparing coins (i.e., between two nickels and a dime). Thus, children have the capacity to view the equal sign in a fully relational way, yet they do not appear to transfer this interpretation to the equal sign when the symbol is presented in arithmetic contexts.

Knuth et al. (2005) also found that children have a poor understanding of the equal sign in the context of arithmetic. Students in Grades 6 to 8 were given the expression "3 + 4 = 7" and asked to explain what the equal sign meant. Even among the Grade 8 children, only 46% gave a relational response. Similarly, Sherman and Bisanz (2004) found an overwhelming tendency among young children to give operator responses when asked what the equal sign meant in symbolic arithmetic contexts. In fact, 94% of Grade 2 children's responses were consistent with an operator interpretation. Such a limited understanding of the equal sign appears to negatively impact performance on arithmetic equivalence problems.

Knuth et al. (2005) explored the relation between students' definition of the equal sign and their performance on equivalence problems. Students in Grades 6 though 8 saw "3 + 4 = 7" and were asked what the name of the symbol was, what it meant, and if the symbol could mean anything else. Responses were coded as being relational, operational, or other. They were then shown "2 x ☐ + 15 = 31 and 2 x ☐ + 15 – 9 = 31 - 9" and asked whether the number that goes in the box was the same for both equations, and then asked students to explain their answers (see description in Chapter I for additional detail). Students who provided a relational interpretation of the equal sign were more likely to realize that the two equations had the same solutions without having to solve and compare than were children who gave an operator interpretation. Thus, how children view the equal sign appears directly related to their approach for problems where there are terms on both sides of the equal sign. An operator interpretation of the equal sign seems to limit children's ability to think about equivalence problems as such. More generally, a misunderstanding of the equal sign has been implicated as a major contributor to difficulties in higher mathematics (Kieran, 1981; Knuth et al., 2005; McNeil & Alibali, 2005a, 2005b).

Evidence of Early Equivalence Skills

Mix's (1999) work with preschool children suggests that children enter school with the ability to detect equivalence across sets of objects or actions. She asked three- to five-year-old children to match sets of objects or events to target sets, and demonstrated that long before formal instruction young children understand equivalence in some conditions. Although Mix's study provides convincing evidence that equivalence in a general sense develops early, what remains unclear is how this early concept of equality with numerical sets relates to later concepts of equivalence in arithmetic. Mix examined children's ability to match or produce equal sets, but the problems did not require any arithmetic operations, unlike those used to test equivalence in older children. This form of *number equivalence* differs from *arithmetic equivalence*, however, in that the latter is reflected in children's ability to judge or make sets of numbers equal by using arithmetic operations. Studies have been done to test children's ability to detect numerical changes to sets after an arithmetic operation has occurred (Gelman, 1972), but the ability to detect arithmetic changes to sets may be different than (and occur prior to) the ability to use arithmetic to create equality on two sides of an expression, called *arithmetic equivalence*. To date, arithmetic equivalence has only been tested in school-aged children using conventional symbols. Thus, we still need to examine the emergence of arithmetic equivalence and whether formal instruction about the equal sign is mapped on to existing concepts of equivalence or if both the symbol and the concept are new to elementary students.

To explore whether children can solve equivalence problems before formal instruction, Sherman and Bisanz (2004) presented kindergarten children with equivalence problems using a novel presentation method that did not contain any conventional mathematical symbols (i.e., *nonsymbolic* problems). Both accuracy and verbal justifications were used to determine a child's solving strategies. Justifications were classified as representing either an equivalence or non-equivalence approach, a method used in previous equivalence research (Alibali, 1999; Perry et al., 1988). Equivalence justifications included all verbal reports whereby children stated that both sides of the expression looked the same or had the same quantity, whereas non-equivalence justifications included responses indicating that children were using some strategy unrelated to ensuring both sides were equivalent, such as the add-all strategy. The authors used three pieces of evidence to determine whether children were thinking about the equivalence problems relationally: (a) accuracy, (b) justifications (equivalent versus non-equivalent), and (c) accuracy and justifications combined on the study's most difficult type of equivalence problem (part-whole problems, such as 5 + 3 = 2 + __).

Overall, Sherman and Bisanz (2004) found that some kindergarten children are indeed capable of using arithmetic to make two sides of an expression equivalent, particularly when the nonsymbolic problems are small in magnitude, as evident by their accuracy, verbal justifications, and combined performance (accuracy and justifications together) on the difficult part-whole equivalence problems. In fact, 24% of the kindergarten children solved all four part-whole problems correctly and gave equivalence justifications, which met the most stringent criterion for determining children's ability to reason about arithmetic expressions relationally. This finding was the first direct evidence that some children enter school with the ability to use arithmetic to make two sides of an expression equivalent. Unfortunately, children's accuracy could not be compared to chance levels because children could respond from an unrestricted range of possible answers. Considering the variable, and often low, accuracy rates, it was difficult to determine whether the children's performance could be considered evidence of understanding the concept of equivalence relationally.

In Study 1 I wanted to build on the previous study by asking kindergarten children to solve nonsymbolic arithmetic equivalence problems using a multiple-choice paradigm so that accuracy could be compared to chance levels. More importantly, I wanted to establish whether children possess an understanding of relational equivalence *before* they formally learn about the equal sign. In Canada, Grade 1 is the year, according to curriculum guidelines, in which children are formally introduced to the symbols and concepts of balance and equality ("=") or imbalance and inequality ("<" and ">") in number sets (Western and Northern Canadian Protocols for Collaboration in Education, 2006). If kindergarten children do hold a relational understanding of equivalence, then it would appear that older children's difficulty with equivalence problems arises from failure to map this understanding onto the symbol (=) as it is taught and used with conventional symbols in mathematics. If they do not, then perhaps more direct instruction and practice with the equal sign, and/or using arithmetic in nonsymbolic contexts, is necessary to teach children about the symbol's significance in mathematics.

Mapping Understanding of Equivalence to the Equal Sign

The nonsymbolic condition is useful for determining whether such young children can solve arithmetic equivalence problems at all. It is also useful for examining how slightly older children may or may not transfer their understanding of equivalence between problems presented both with and without symbols. In previous research (Sherman & Bisanz, 2004), we found convincing evidence, from both accuracy and verbal justification data, to suggest that Grade 2 children can solve problems when they are presented without conventional symbols, yet other children, when given the same problems with conventional symbols, overwhelmingly fail. Therefore, in Study 2, I wanted to extend the previous research in two important ways: (a) a multiple-choice format was used so that accuracy in both conditions could be compared to chance levels, and (b) a within-subject design was used so I could assess potential order effects. The first modification is important for determining with confidence whether information about accuracy provides evidence for or against the assertion that young children understand equivalence relationally in the context of arithmetic.

The second modification is essential in exploring the issue of mapping. Specifically, Grade 2 children's undeniable overall success on nonsymbolic problems in the previous study (Sherman & Bisanz, 2004) raises the question of whether practice with equivalence problems in this condition might be useful for learning about arithmetic relationally and subsequently solving symbolic equivalence problems successfully.

Children may map their strategies for making both sides of the nonsymbolic equations equal to symbolic problems. In light of the mediocre success that other training and intervention studies have had on arithmetic equivalence performance (Alibali, 1999; McNeil & Alibali, 2000; Perry, 1991), the mapping hypothesis has an appealing practical application: Simply solving equivalence problems in the nonsymbolic condition might help performance on symbolic problems.

Study 1

In previous work Sherman and Bisanz (2004) provided evidence to support the argument some kindergarten children reason about equivalence problems relationally, as evident by their accuracy and verbal justifications. Using the nonsymbolic condition allowed us to test children who had not received any formal instruction about mathematics (or about mathematical symbols) and to determine whether these children enter school with the ability to make two sides of a mathematical expression, in which both sides contain more than one addend, equal. Examining the concepts and skills with which children enter school is important for understanding cognitive development in general, but also for understanding how formal instruction and prior knowledge interface more specifically. That is, evidence that kindergarten children solve equivalence problems successfully would indicate that children enter school with the ability to use arithmetic to solve these complex problems, but that this ability does not get mapped onto typical, symbolic problems later on. Unfortunately, children's accuracy in the previous study (Sherman & Bisanz, 2004) was difficult to interpret in isolation because performance could not be compared to accuracy that could be expected with random guessing. Therefore, in this study I tested kindergarten children with nonsymbolic problems and with answer options to assess their ability to solve equivalence problems, evaluate the type of solutions they used, and compare their accuracy to chance levels.

I hypothesized that children will show some evidence of the ability to solve arithmetic equivalence problems (i.e., mean performance will be higher than expected by chance), but that performance will be, on average, quite low due to the high arithmetic demands. I also hypothesize, however, that some children will show convincing evidence of understanding equivalence, as indicated by high levels of choosing the relational responses, based on the high accuracy by some of the children in the previous study (Sherman & Bisanz, 2004). In this and the following studies, I examined children's accuracy in relation to gender, as gender sometimes has been found to be related to mathematical performance (Felson & Trudeau, 1991; Hyde, Fennema, & Lamon, 1990), mathematical strategy use mediated by temperament (Davis & Carr, 2002), and attitudes toward mathematics (Hyde, Fennema, Ryan, Frost, & Hopp, 1990) to varying degrees throughout the school years.

In addition to accuracy, I examined children's justifications to determine whether and how often they represent the same strategies as evident by their corresponding solution options. If the solution option and coded justification reflected the same strategy (e.g., an add-all solution option and a justification coded as being add-all), then children were considered to have a *match*, or agreement, for that particular problem. Matches provide convincing evidence that children deliberately chose the solutions and could explain why they did so. Analyzing matches and mismatches helped me address four questions: (a) are children more likely to have matches for operator compared to relational strategies, (b) what strategies are children using, (c) does strategy use relate to accuracy levels, and (d) when children do choose the relational solutions, how often are they also giving a relational justification as opposed to mismatching justifications? I hypothesized that children will have high levels of mismatches, given that the young children may have difficulty verbally describing their strategies, such as adding all terms to the left of the equal sign, particularly since they have not received formal training in the conventional procedures or terminology of arithmetic. Furthermore, I hypothesize that more mismatches will occur for relational solutions, as explaining that both sides are equal may be more difficult than describing an operator procedure such as adding all numbers together. I hypothesized that some children will solve the problems relationally, whereas others will use mainly operator or guessing approaches, and that strategy use will relate to accuracy.

Method

Participants and Design

Participants consisted of 31 kindergarten children, including 17 males and 14 females, from two public schools. Ages ranged (in years; months) from 4;8 to 6;6, with a mean of 5;8. Three additional children were tested but not included in analyses because they did not finish the session. Because the focus of this and subsequent studies was on children's changing performance on equivalence problems, rather than correlates of individual differences, student characteristics, including reading and general mathematical skill levels, were not assessed. Because there is some evidence to suggest that language skills may relate to children's ability to learn aspects of mathematics, specifically algebraic notation (e.g., MacGregor & Price, 1999), including language-related measures in future studies would be useful.

Materials and Procedures

Practice problems. Kindergarten children participated in one 20-minute session that began with five practice problems to familiarize them with the materials. To present all problems, wooden blocks approximately 5 cm high were placed into opaque plastic bins approximately 3 cm high to represent each set (or number) in the equations. A 6-cm high piece of blue cardboard, folded to look like a tent, served to separate the two sides of the equations (Appendix A). The two sides of the equation were made more salient by placing purple construction paper under the bins on one side of the blue tent and orange paper on the other side. The practice problems included a counting string (1, 2, 3, 4, __, 6), two basic arithmetic problems (2 + 1 = __ and 1 + 4 = __, with 2 + 2 = __ as a back-up problem), and three equivalence problems (2 + 2 = 1 + __, 1 + 2 + 1 = 1 + __, and 1+1=1+__). On the counting string, the first arithmetic problem, and the first two equivalence problems, children were asked to produce solutions by placing blocks into the empty bin. On all problems, except for the counting string, children were instructed to "put blocks (or a photograph of the blocks) into the empty bin so that when you put together these [experimenter points to the blocks on the right side] on this (right) side of the blue tent, you'll have the same number as when you put together these [experimenter points to the blocks on the left side of the tent] on this (left) side of the blue tent."

After solving the problems the solution blocks were removed from the bin and children were asked to solve the same question again by choosing the correct answer from a choice of four photographs (Appendix A). The photographs were of various quantities of blocks within a bin, each of which corresponded to four specific ways of approaching the equivalence problems (add-all, add-to-equal, relational, and choosing the smallest quantity). Add-all, add-to-equal, and relational approaches have been documented in many studies (Alibali, 1999; Goldin-Meadow, Kim, & Singer, 1999; McNeil & Alibali, 2000; Perry et al., 1988; Rittle-Johnson & Alibali, 1999). The small-number choice was developed for this study because children may tire of counting and choose the smallest solution to save from having to count higher. Without a small-number option, children who decided on the smallest solution to avoid excess counting would by definition be choosing the relational answers (which are always smaller than adding all terms or adding terms to the left of the equal sign). This approach would arbitrarily inflate the proportion of children who appeared to solve the problems relationally. Along with the four photographs there was an additional option of a question mark for problems in which children thought none of the solutions were correct. To ensure that children considered all five choices when solving all problems, one of the practice problems required that children use the question mark to answer correctly. After solving each practice problem, feedback was given and children were asked to justify their response by telling the experimenter why his or her answer was correct. Justifications were coded using the criteria described in Appendix B.

Test problems. Children received 10 problems in the nonsymbolic context. To assess whether children performed differently on specific types of problems, half of the problems were of the *combination* type ($a + b + c = a + \_\_$), the other half were of the *part-whole* type ($a + b = c + \_\_$) (Appendix C). Both problem types have been found in previous research to be challenging equivalence problems (Sherman & Bisanz, 2004) and similar to the form used in other equivalence research (Alibali, 1999; Goldin-Meadow et al., 1999; McNeil & Alibali, 2000, 2002, 2004, 2005b; Perry, 1991; Perry et al., 1988).

There were never more than two of the same problem types in a row. For all of the equivalence problems, children solved using the photographs. Instructions were given for the first two problems of the problem set, and, following each solution, children were asked to justify their responses. Solutions, any observable solving behaviors, and verbal justifications were recorded (Appendix B), and no feedback was given.

Arithmetic problems. Completion of the problem set was followed by six canonical arithmetic problems (4 + 3 = __, 6 + 4 = __, 1 + 3 = __, 2 + 7 = __, 3 + 5 = __, and 2 + 3 = __), presented in the nonsymbolic condition. Children solved the problems by placing blocks in the empty (last) bin. The problems were within the same magnitude range as the equivalence problems to ensure that children were capable of the arithmetic necessary to solve the equivalence problems. High accuracy on arithmetic problems would circumvent the argument that failure on equivalence problems was due to poor arithmetic skills in general. No feedback was given.

Results

Overall Performance

I looked at overall performance in two ways. First, I compared the mean accuracy of all kindergarten children on the equivalence problems to the level of success expected by guessing to determine whether, overall, kindergarten children were showing evidence of correctly solving equivalence problems. Second, I used an ANOVA to explore potential effects of gender and problem type. The probability of responding correctly by guessing was considered to be 0.20 (or 20%) because there were five solution options given with each equivalence problem. The children's overall mean was fairly low (*M* = 33.5%, *SD* = 37.6%), with a 95% confidence interval from 20.3% to 46.7%. Thus, the chance level of accuracy (two out of ten) is slightly below the likely range of the true mean as estimated using the sample mean and standard error.

Although the mean accuracy level was fairly low, performance on the equivalence problems did not appear to be due to low general arithmetic skills. Kindergarten children had fairly high accuracy ($M = 75.8\%$, $SD = 34.7\%$) on the six arithmetic problems presented after the equivalence problem set, and there was no correlation between scores on the arithmetic problems and accuracy on the equivalence problems, $r(29) = .07$, $p = .69$. Thus, the relatively low overall performance on equivalence problems cannot be attributed to constraints on addition skill. The criticism can be made that the equivalence problems demanded higher levels of arithmetic skill than the arithmetic problems, given the difference in problem magnitude and number of terms. The lack of correlation between the two tasks, however, is consistent with the view that children's performance on equivalence was unrelated to simple arithmetic skill.

To determine whether accuracy differed across the two types of problems, as well as across gender, a 2(Gender) x 2(Problem type: part-whole, combination) ANOVA was conducted with repeated measures on the last variable, even though accuracy was distributed non-normally (Figure 2-1). Accuracy on part-whole problems ($M = 37.4\%$, $SD = 39.2\%$) was marginally but significantly higher than accuracy on combination problems ($M = 29\%$, $SD = 38.2\%$), $F(1, 29) = 6.68$, $p < .05$. No other main effects or interactions were significant, and gender will not be discussed further. Perhaps the fact that part-whole problems had one less term to add (i.e., one less bin with blocks) influenced performance, despite the fact that with the part-whole problem type children *must* use arithmetic because none of the terms are the same on both sides of the blue tent.

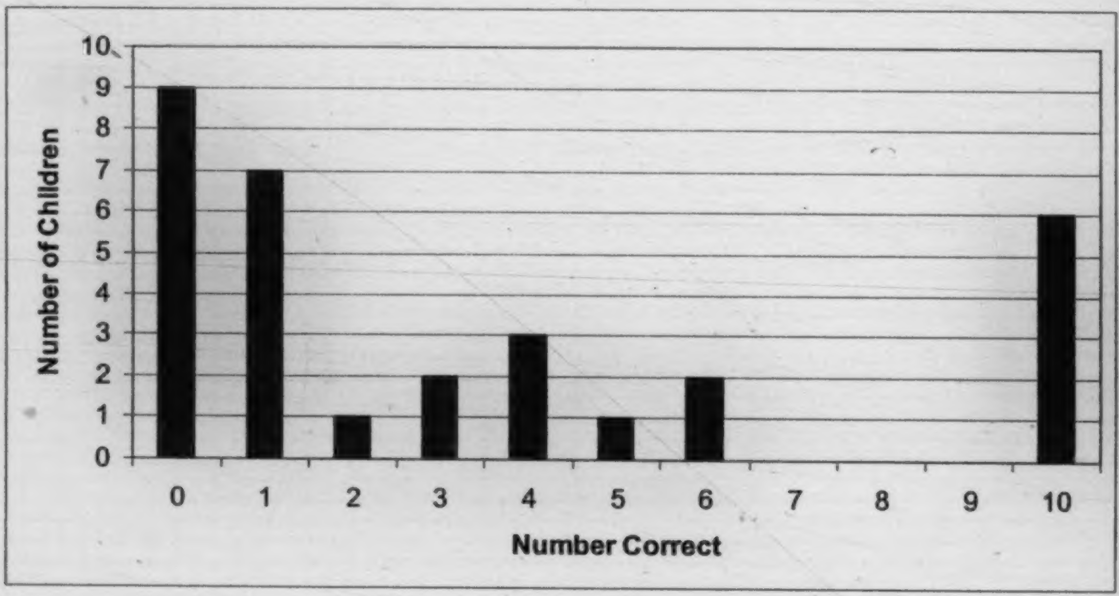
Perhaps prior to formal mathematic instruction and experience children are more influenced by the number of terms to calculate and hold in working memory than by the need to use arithmetic per se (Klein & Bisanz, 2000; Rasmussen & Bisanz, 2003).

Individual Differences

Several children demonstrated surprising competence, considering their young age, in using arithmetic to solve the problems relationally. Although the effect of problem type was significant, the most striking part about children's responses was their variability and the individual differences they demonstrated with respect to accuracy (Figure 2-1) and strategies. The distribution of children's accuracy on the equivalence problems prompted examination of differences between children who invented a correct, relational strategy versus those who were less successful. I first discuss the accuracy and strategies of the children who had extremely high accuracy on the ten equivalence problems, followed by those who solved no equivalence problems correctly. Finally, I examined the group of children whose accuracy fell in the middle range.

I wanted to explore the strategies that children used to solve these novel, challenging problems in an attempt to determine whether children's approaches fit into comprehensive patterns. For example, do children with both very high and very low accuracy use strategies (whether correct or not) more consistently then children whose accuracy fell in the middle of the range? Are children with low to medium accuracy simply inventing inappropriate strategies, guessing randomly, or using a variety of approaches? Do children's justifications match their solution options, and what do mismatches between solutions and justifications mean?

Figure 2-1. Kindergarten children's accuracy on 10 arithmetic equivalence problems presented nonsymbolically.

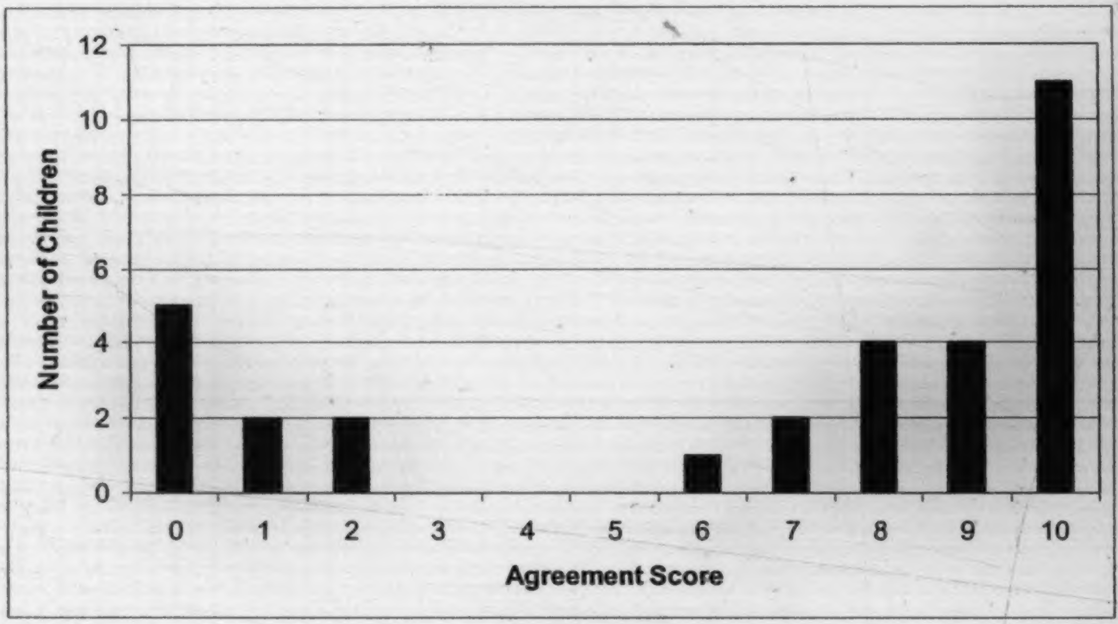


High-accuracy group. The binomial probability of correctly solving five or more problems by guessing, with an α of .05 and the probability of guessing correctly assumed to be .2, is .03. In fact, nine children (29%) solved five or more problems correctly, six of whom (19%) were correct on all 10 problems (Figure 1). Thus, a sizable number of children solved more problems correctly than expected by chance.

Overall, children's justifications matched their solution choices on 65.8% of all equivalence problems, although there was a wide range in the number of matches across all participants (Figure 2-2). For children in the high accuracy group, justifications matched solution choices on 86.7% of the problems. In fact, the six children who solved the ten equivalence problems correctly had no mismatches. That is, these children gave relational justifications to support their relational choices for all 10 problems. Among the remaining three children in the high accuracy group, one child had 10 matches, one had seven, and another had one. The child with 10 matches told the experimenter that there were two ways to solve the problems. She began by solving the first four problems using an add-to-equal approach and justifying her responses accordingly. On the fifth problem, however, she switched to a relational approach and gave relational justifications for the remaining six problems. The child with seven matches solved six problems correctly, providing relational justifications on five problems but an ambiguous justification for the sixth correct problem. He also solved two problems with an add-to-equal strategy, supported by matching justifications. He chose the small number solution on two problems, with his corresponding justifications being classified as ambiguous in one case and pattern/matching in the other case. Finally, the child with one match gave ambiguous justifications on eight problems. That is, she chose several different solution options, including relational, add-all and small number, but her justifications were not interpretable. She matched solution and justification for one problem, solving in an add-all manner.

Overall, children in the high accuracy group can be characterized as choosing the relational solutions more often than can be expected by random guessing, and, in most cases, supporting these correct choices with relational justifications. For a large proportion of this group, children invented a relational strategy that enabled them to achieve perfect accuracy. For two children, justifying solutions ambiguously lowered the degree of agreement between solution choices and justifications, and one child clearly attempted to solve the equivalence problems in two distinct ways: add-to-equal and then a relational strategy. Given their young age, it is not surprising that some kindergarten children found it difficult to verbally explain why they chose their solutions, yet only two children from this group gave ambiguous justifications.

Figure 2-2. Distribution of matches between solution choices and justifications across the ten equivalence problems for all 31 kindergarten children. Matches were scored as one point, and mismatches zero, for a maximum agreement score of 10.



Zero-accuracy group. The binomial probability of getting none of the 10 equivalence problems correct by guessing, with an α of .05 and the probability of guessing correctly being .2, is .11. Remarkably, nine children (29%) were incorrect on all 10 equivalence problems (Figure 2-1). Solutions matched justifications for 73.3% of all equivalence problems. Four of the nine students had matches for all 10 problems, two matched on nine problems, and one matched on eight problems. All seven of these children used add-all and/or add-to-equal strategies, as evident by their matching solutions and justifications. Of the two children with no agreement between solutions and justifications, one had ambiguous justifications on all 10 problems (and chose add-all solutions), the other child gave pattern/matching justifications and mainly (9 of 10 problems) chose the question mark solution.

Overall, children who solved all ten problems incorrectly invented operator (i.e., add-all and/or add-to-equal) strategies for solving, as evident by the high rate of agreement between their solutions and justifications. Children who solved the equivalence problems either very successfully or completely wrong did so with high agreement between solutions and justifications. High agreement is useful for determining that children were deliberately using a particular strategy to solve the equivalence problems. What makes these two groups interesting is the consistency with which they approached the problems. Using problems where children's solutions and justifications matched, only three of the 18 children used more than one strategy. Two of these children, both from the high accuracy group, used add-to-equal and relational strategies, whereas the other child, from the zero accuracy group, used both add-all and add-to-equal strategies. Of the eleven children who had perfect agreement scores, all were from either the zero accuracy (four) or high accuracy (seven) groups.

Low-accuracy group. Thirteen children (41.9%) solved between one and four equivalence problems correctly (Figure 2-1). In addition to their low accuracy, children in this group had fewer matches between their solutions and justifications (46.2%) than children from the other two groups. Because fewer than half their solutions were supported by corresponding justifications, it was challenging to characterize how children from this group were approaching the equivalence problems. Examining solution choices alone may not capture the true range of approaches children were using, and may reflect guessing as opposed to the deliberate invention or use of a strategy to solve the problems. Similarly, examining justifications alone may not accurately capture children's strategy use, as children often gave ambiguous justifications, or justifications that did not support the corresponding solution choices. Indeed, mismatches may be evidence of children guessing, or evidence of them trying multiple approaches, perhaps within a single problem. That is, by selecting a solution option that represents one approach, while justifying a different approach, children may have been attempting to increase their chances of looking or sounding correct.

Because of the high degree of disagreement between solutions and justifications, and because examining only the solution choices or justifications alone had potentially important limitations, I examined just the cases where children's actions and statements did match. Matches were considered to be convincing evidence of strategy use. Three questions were addressed: (a) what strategies were children from this low-accuracy group using, (b) were children more likely to have matches for operator compared to relational strategies, and (c) when children did choose the relational solutions, how often were they also giving a relational justification as opposed to mismatching justifications?

Children's solutions and justifications matched only for relational, add-all, and add-to-equal strategies. Of the 130 total responses across all ten problems by children from this group, 10% were matched relational strategies, 24.6% were matched add-all strategies, and 11.5% were matched add-to-equal strategies. Thus, 36.2% of these children's responses were clearly and convincingly (i.e., solutions and justifications matched) operator strategies. When children did chose relational solutions, their justifications mismatched more often than matched. That is, there were a total of 29 relational solutions across the equivalence problems by the 13 children in this group, with every child choosing the relational solution on at least one occasion, yet only 13 matching relational justifications. The mismatching justifications took one of two forms: Children either gave ambiguous justifications (on 5 occasions) or pattern/matching justifications (on 11 occasions). The large number of mismatches may or may not be indicative of children guessing when they chose the relational responses. For instance, ambiguous justifications might be a sign that children are simply unable to adequately explain why they thought their solution was correct, rather than an indicator that they do not know why their answer is correct. Similarly, pattern/matching justifications highlight the fact that children are attending to the visual similarities between two sides of an equation, which may be a rudimentary component of using arithmetic to ensure *numerical* similarity between two sides of an equation. Regardless of precisely what mismatches between relational solutions and justifications mean, children in this low accuracy group showed convincing evidence of using a relational strategies on very few problems. Relational solutions are supported by relational justifications on less than half of the problems, and matched relational strategies make up a very small proportion (10%) of all responses by children in this group.

There were two children whose strategies highlight the variability and creativity evident in many of the kindergarten children. One child only solved one of the 10 problems correctly and chose the add-to-equal responses consistently. His justifications and gestures indicated, however, that he was attempting to make both "sides" equivalent but that he failed to include the term to the right of the blue tent. Thus, the strategy he invented to solve these complex arithmetic problems was to deliberately ignore the term to the right of the blue tent, to ensure equality between the left side and the empty bin, and to verbally refer to his approach as making both "sides" the same.

Another child, who solved four equivalence problems correctly, demonstrated how vulnerable her relational strategy was. She chose the relational solution and gave a relational justification only on problems where I actually saw her deliberately attend to the *c*/*d* term. Whenever I noticed that she did not obviously look directly at the *c*/*d* term, she would chose the add-to-equal solution. Unlike the other child who indicated attempts to make both "sides" the same even while adding-to-equal, this child would give add-to-equal justifications to support her add-to-equal solution. Therefore, unless she obviously noted the *c*/*d* term while trying to solve the problem, she ignored it in both calculation and justification. Her ability to solve relationally depended on attending to the term to the right of the blue tent.

I also examined children's solutions to determine whether there where instances whereby children gave relational justifications but did not chose relational solutions. One child from this low-accuracy group gave relational justifications for all 10 problems but only chose the relational solution once. On the majority (seven) of the problems, he chose the add-to-equal solutions, and the experimenter noted that the child failed to attend to the *c/d* term to the right of the blue tent. By his justifications it is clear that he believed he was making both sides of the tent the same even though he was in fact attending to just a portion of the equivalence problem. He chose both the small number and question mark option for one problem each. There was one other child who gave a relational justification but chose a mismatching (small number) solution on just one problem. Because she chose relational solutions every other time she gave a relational justification, her small number choice likely represents a calculation rather than conceptual error. In no other cases did children five relational justifications but not choosing the corresponding relational solutions.

Discussion

By presenting problems nonsymbolically, I was able to assess young children's understanding of complex problems in the absence of conventional mathematical symbols. Overall, some kindergarten children clearly demonstrated an understanding of relational equivalence in the nonsymbolic format that, according to standard curriculum guidelines, precedes formal instruction about equivalence in arithmetic contexts. The multiple-choice paradigm and analysis of children's justifications allowed me to confidently state that more than a quarter of the children tested showed evidence of solving equivalence problems relationally. The overall mean performance level was higher than would be expected by random guessing, but what the overall mean did not indicate was the variability across and within children, as well as the surprisingly precocious skills of many of the children. Almost 30% of the children solved more than half of the equivalence problems correctly, and most of these children solved all 10 problems accurately. Equally interesting was the finding that another 30% of the children solved all 10 equivalence problems incorrectly. Children in both the high- and zero-accuracy groups tended to have a high degree of agreement between the solutions they chose and the justifications they gave. Matches between these two types of data helped characterize their strategies such that some children clearly used a relational strategy to solve the problems, whereas other children clearly invented operator strategies (i.e., add-all and add-to-equal) that are common among older children. All 11 children who had perfect agreement scores for solutions with matching justifications came from either the zero- or high-accuracy groups.

Approximately 40% of the children were in a low accuracy group, solving between one and four equivalence problems correctly. Children from this group can be characterized as having high levels of disagreement between their solutions and justifications, which may indicate higher rates of guessing. When there were matches between solutions and justifications, the matches were more likely to be for operator as opposed to relational strategies. The high number of operator strategies is interesting in the sense that practice with canonical arithmetic problems and exposure to other left-to-right activities, such as reading, is often associated with adherence to operator patterns and strategies (McNeil & Alibali, 2005b; Seo & Ginsburg, 2003). Kindergarten children have not been exposed to much formal mathematical education, but rather appear to be inventing these approaches as potential methods for solving equivalence problems. Results from this study indicate that some children begin formal schooling with the ability to invent an accurate, relational strategy, while others begin with the capacity for operator strategies.

Study 2

Although it appears that some children enter school with a relational approach to complex arithmetic problems, children in elementary school overwhelmingly fail such problems in the conventional symbolic context. In a previous study, Grade 2 children who solved nonsymbolic problems had very high accuracy, but their Grade 2 peers who solved the same problems in the symbolic context had very low accuracy (Sherman & Bisanz, 2004). Taken together, it appears as though children are capable of reasoning relationally in complex arithmetic problems, yet they may not map their knowledge of equivalence onto problems presented with conventional symbols, perhaps due to their limited, operator view of the equal sign. In this study, I examined whether presenting problems in *both* contexts (nonsymbolic and symbolic) may help the mapping process and thus improve performance on symbolic problems. By solving the same problems in both contexts, children may recognize the similarity between the problems or the task and instructions, and realize that thinking relationally in both contexts will lead to success. Specifically, I hypothesized that children who solve nonsymbolic problems first will have high accuracy in the nonsymbolic problems and will continue to perform with high accuracy on subsequent symbolic problems, mapping their understanding of equivalence onto the equal sign in the symbolic context. I assessed this by examining accuracy on symbolic equivalence problems and by specifically asking students to define the equal sign.

Patterns of solution choices were also examined to determine whether children's strategies for solving equivalence problems, not just their accuracy, differed in relation to the order in which they received the sessions. I hypothesized that children who solved nonsymbolic problems first would use a relational strategy more than other strategies on both sessions, whereas children who began with symbolic problems would use primarily operator strategies in the symbolic session and the relational strategy in the nonsymbolic session.

I hypothesized that if children's performance on symbolic problems improves after experience with nonsymbolic problems, then their verbal interpretations of the equal sign might be relational in nature also. That is, experience with nonsymbolic problems may help children map a relational understanding of the equal sign onto the symbol itself. I also anticipated that children will have low accuracy on the symbolic problems, as found in other studies using symbolic equivalence tests, but that children will improve their performance on nonsymbolic problems. Children will likely approach the nonsymbolic problems relationally and solve with high accuracy, based on the high performance demonstrated in the previous study with Grade 2 children (Sherman & Bisanz, 2004).

Method

Participants and Design

Participants consisted of 32 Grade 2 students, including 16 males and 16 females, from four local public schools. Ages ranged (in years; months) from 7;2 to 8;9, with a mean of 7;9. One additional female was tested but not included because she did not solve any of the arithmetic problems correctly in the nonsymbolic condition. Because low accuracy on equivalence problems could be attributed to poor arithmetic skill, and because solving the equivalence problems may have been stressful for her if she did indeed have difficulty with simple arithmetic, this particular student was excused from the study. All children participated in two 20-minute sessions, with approximately one week between sessions. Both sessions began with a variety of practice problems used to familiarize them with the materials and instructions. Each child was tested by the same experimenter for both sessions, except for one Grade 2 child who was tested by a different experimenter for each session.

To assess whether performance differed across the two conditions, children received the same 16 problems in both the symbolic and nonsymbolic contexts. To determine whether children's performance was affected by order of experience with the conditions, children were randomly assigned to one of two session orders: symbolic then nonsymbolic (symbolic/nonsymbolic, including 8 males and 8 females) or nonsymbolic then symbolic (nonsymbolic/symbolic, including 8 males and 8 females).

Materials and Procedures

Nonsymbolic condition. All problems were presented using the same materials as in Study 1, and the instructions were identical. The practice problems included a counting string (1, 2, 3, 4, __, 6), one basic arithmetic problem (2 + 1 = __, with 2 + 2 = __ as a back-up problem), and a practice equivalence problem (1 + 1 = 1 + __). Grade 2 children received one less practice problem than the kindergarten children, as I believed they would require less exposure to the problems to understand the task and instructions compared to the younger children in Study 1. Feedback was given for all practice problems and children were asked to justify their responses (Appendix B).

Following the practice problems, the problem set (consisting of 16 equivalence problems, Appendix C) was administered. As in Study 1, instructions were given for the first two problems and students were asked to justify their responses. Solutions, any observable solving behaviors, and verbal justifications were recorded. No feedback was given. Children solved the problems by choosing among four photographs of blocks that corresponded to four methods of solving equivalence problems, with the fifth choice being the question mark. Children were instructed to use this card if they believed none of the choice cards were correct. The question mark was the correct choice for one of the practice problems, but none of the equivalence problems in the problem set. In the problem set, the order in which the choice cards were presented was determined using a Latin Square for the first and second half (8 problems each) of the problem set. Order of choice cards was never the same for two questions in a row. As in Study 1, half the problems were part-whole types and half were combination types. There were never more than two of the same problem types in a row.

If children chose the same type of solution option (e.g., add-all) on at least 13 of the 16 equivalence problems (approximately 80%), they were classified as consistently solving in that manner. For example, if a child chose the add-all solution choice on 13 or more problems, he or she was classified as using an add-all strategy. The binomial probability of randomly selecting the same type of solution option on 13 of 16 problems, with an α of .05 and the probability of choosing any one type of solution option being 0.2 (one of five options), is less than .0001.

Symbolic condition. All symbolic problems were presented using a black stand placed on top of a table or desk. White laminated cards containing the problems were flipped over one at a time. Prior to the presentation of each problem, solution choice cards were arranged along a horizontal line beside the problem card, fixed with Velcro (Appendix A). Each problem contained a blank line and Velcro strip to indicate where the solutions needed to be placed. Children solved by removing the solution card of their choice from the line and fixing it to the blank line on the problem card. As in the nonsymbolic condition in both Studies 1 and 2, choice cards corresponded to add-all, add-to-equal, relational, and small number approaches. An additional choice card, depicting a question mark, remained available for all problems. Instructions were given for the first two problems, and students were asked to justify their responses. Solutions, any observable solving behaviors, and verbal justifications were recorded, and no feedback was given. For all equivalence problems, verbal justifications were recorded in case children's solving strategies could not be interpreted based on their solutions (Appendix B). As described above, children were credited with using an approach consistently if they solved in the same manner on 13 or more problems.

For both the symbolic and nonsymbolic sessions, scores for the problem set were converted to percent correct because of experimenter error. That is, 12 children received only 15 equivalence problems due to a duplicate problem in the problem set. Once the duplicate was detected, a new problem was created for the remaining participants.

Arithmetic problems. Completion of the problem sets for both sessions was followed by presentation of six arithmetic problems (4 + 3 = __, 6 + 4 = __, 1 + 3 = __, 2 + 7 = __, 3 + 5 = __, and 2 + 3 = __), presented in the same manner as the problem set. Nonsymbolic problems were solved using blocks, whereas symbolic problems were solved by children writing their solutions on the problem cards using a dry-erase pen.

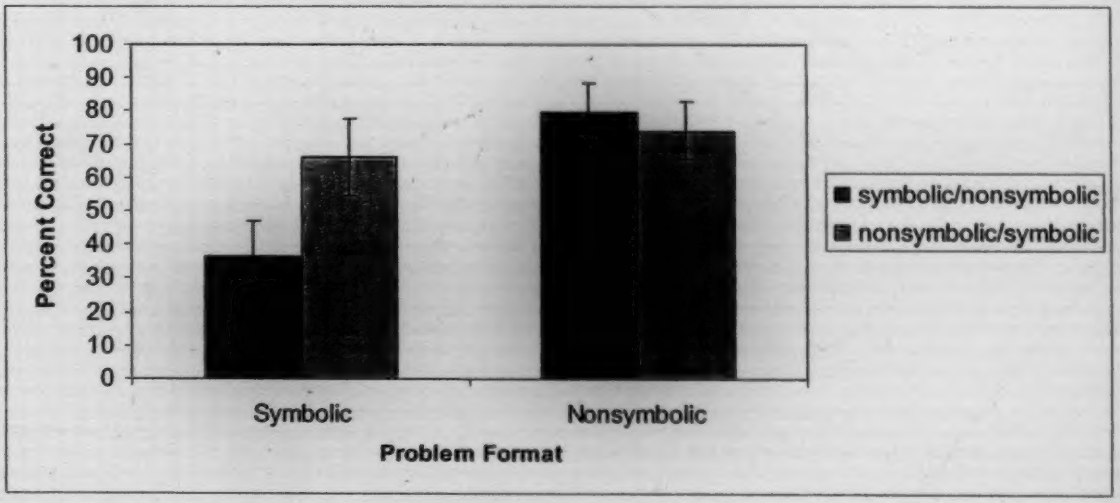
The problems were within the same magnitude range as the equivalence problems, and were designed to assess addition skill level so that, should children perform poorly on equivalence problems, failure could not be attributed to poor general arithmetic skills. No feedback was given.

Verbal interpretation of the equal sign. Following completion of the arithmetic problems in the symbolic context, children were asked what the equal sign means. Specifically, the experimenter would cover both sides of the last symbolic arithmetic problem (2 + 3 = __), leaving just the equal sign exposed, and ask "what is this?", followed by "what does it mean?" Children's explanations were coded (Appendix D), and no feedback was given. Children were prompted to respond again and "tell me more" if they stated that the equal sign means "equals."

Results

Accuracy was analyzed using a 2(Session order: symbolic/nonsymbolic, nonsymbolic/symbolic) x 2(Gender) x 2(Condition: symbolic, nonsymbolic) x 2(Problem Type: part-whole, combination) ANOVA with repeated measures on the last two variables. Children had much higher accuracy on nonsymbolic problems ($M = 76.9\%$, $SD = 34.8\%$) than symbolic problems ($M = 51.3\%$, $SD = 46.1\%$), $F(1, 28) = 12.08$, $p < .01$, and condition interacted with session order, $F(1, 28) = 5.52$, $p < .05$ (Figure 2-3). In general, children in the symbolic/nonsymbolic group solved few symbolic problems and many nonsymbolic problems correctly. Children in the nonsymbolic/symbolic group, however, had high accuracy on both symbolic and nonsymbolic problems. That is, although accuracy on nonsymbolic problems was high for children in both the symbolic/nonsymbolic group and the nonsymbolic/symbolic group, children who solved

Figure 2-3. Mean accuracy by Grade 2 children (symbolic/nonsymbolic and nonsymbolic/symbolic session order groups) on equivalence problems presented in both symbolic and nonsymbolic conditions, with standard error bars.



nonsymbolic problems first (nonsymbolic/symbolic) had much higher accuracy on symbolic problems than children who solved symbolic problems first (symbolic/nonsymbolic) (Figure 2-3). Solving problems in the nonsymbolic condition facilitated performance on subsequent symbolic problems, even though the sessions were approximately a week apart and experimenters did not deliberately or explicitly draw children's attention to the similarity between the two "tasks" across sessions. Accuracy did not differ between part-whole and combination problems, and there were no other significant main effects or interactions. Scores on the six arithmetic problems were quite high (*M* = 94.8%, *SD* = 18.7% in the symbolic condition; *M* = 95.8%, *SD* = 10.3% in the nonsymbolic condition) and are not discussed further.

The multiple-choice method was useful for two reasons: (a) it allowed me to compare accuracy rates to chance levels to help determine how well children solved equivalence problems, and (b) it enabled me to examine children's solving strategies for consistent patterns (i.e., did children choose the add-all solutions on most trials, were they guessing randomly, etc.). With respect to the first benefit, the probability of guessing correctly was considered to be 20% because of the five solution options given on every trial (add-all, add-to-equal, relational, small number, and other). For children in the symbolic/nonsymbolic group, performance on symbolic problems (*M* = 36.5%, *SD* = 42.3%) was particularly low. A 95% confidence interval indicated that the true mean was between 15.9% and 57.9%, a range that includes the observed accuracy level that could be expected with random guessing. Therefore, for children who solved symbolic problems in their first session, their accuracy on symbolic problems was not significantly higher than could be expected by guessing. The other three means (children in symbolic/nonsymbolic group on nonsymbolic problems, and children in nonsymbolic/symbolic group on both symbolic and nonsymbolic problems) were higher than would be expected by chance (Figure 2-3), with confidence intervals ranging from 43.3% to 97.1%.

In Study 1, kindergarten children showed high variability in their accuracy on nonsymbolic equivalence problems. By grouping children into zero-, low-, and high-accuracy groups, I was able to examine specific characteristics within each group, such as the high levels of agreement between solution choices and justifications among children who either got none or most of the equivalence problems correct. The same analyses were not done in Study 2, however, because Grade 2 children did not show the same amount of variability within each session order group and condition and because there were fewer children to compare. For example, the accuracy of the 16 children in the symbolic/nonsymbolic group on symbolic problems was very low overall, with only two children who solved all 16 problems correctly. Similarly, the accuracy of the 16 children in the nonsymbolic/symbolic group on nonsymbolic was very high overall, with only two children who would be classified as zero-accuracy.

Strategy Use

Having solutions that corresponded to specific ways of thinking about and solving equivalence problems allowed us to examine whether children could be categorized as solving in a certain manner (such as an *add-all* approach to equivalence problems). Overall, children's matches between solutions and justifications across the 32 problems was quite high ($M = 79.7\%$, $SD = 28.8\%$). Because children so often supported their solution choices with their justifications, solution options alone were used to examine strategy use by children in the two session order groups. If children responded in the same fashion on at least 13 of the 16 problems (approximately 80%), they were classified as solving in that manner.

Because children's accuracy on symbolic and nonsymbolic problems was affected by the order in which they received the problems, classifications of strategy use for the two conditions were examined separately for children in the session order groups. Children in the symbolic/nonsymbolic group displayed a variety of strategy classifications when solving symbolic problems (Figure 2-4, left bar). Five children consistently used a relational strategy, three used an operator strategy, three tried to solve by making patterns or matching terms on both sides of the equations, and three children used a variety of approaches or solved part-whole versus combination problems differently (i.e., part-whole problems solved with an add-to-equal strategy, combination problems solved relationally). These same children presented a very different pattern of strategy use when solving the nonsymbolic problems a week later (Figure 2-4, right bar), with twelve of the children using a relational strategy.

Children in the nonsymbolic/symbolic group could be characterized as using one of three strategies to solve the nonsymbolic problems (Figure 2-5, right bar), with the majority solving the problems relationally. When solving symbolic problems a week later, children could be classified as using one of five strategies (Figure 2-5, left bar), but the majority of children (11 of 16) still solved the problems relationally. Relational justifications almost never occurred in the absence of relational solution choices, occurring on fewer than 2% of problems. Therefore, when children said they solved relationally, they almost always chose the relational solution.

Figure 2-4. Number of the 16 Grade 2 children in the symbolic/nonsymbolic group who used specific strategies on both symbolic and nonsymbolic sessions.

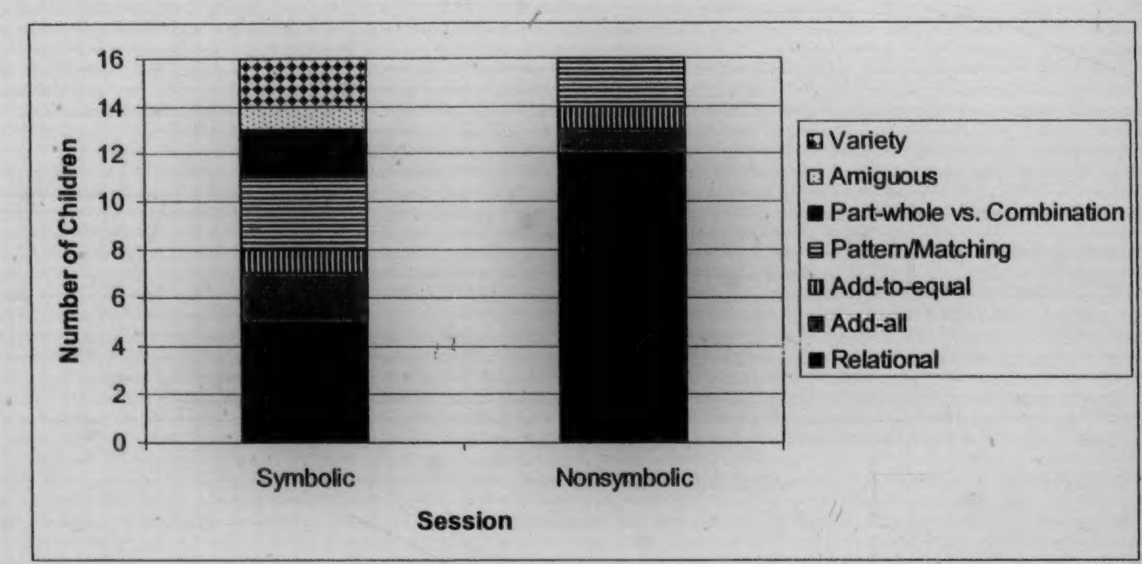
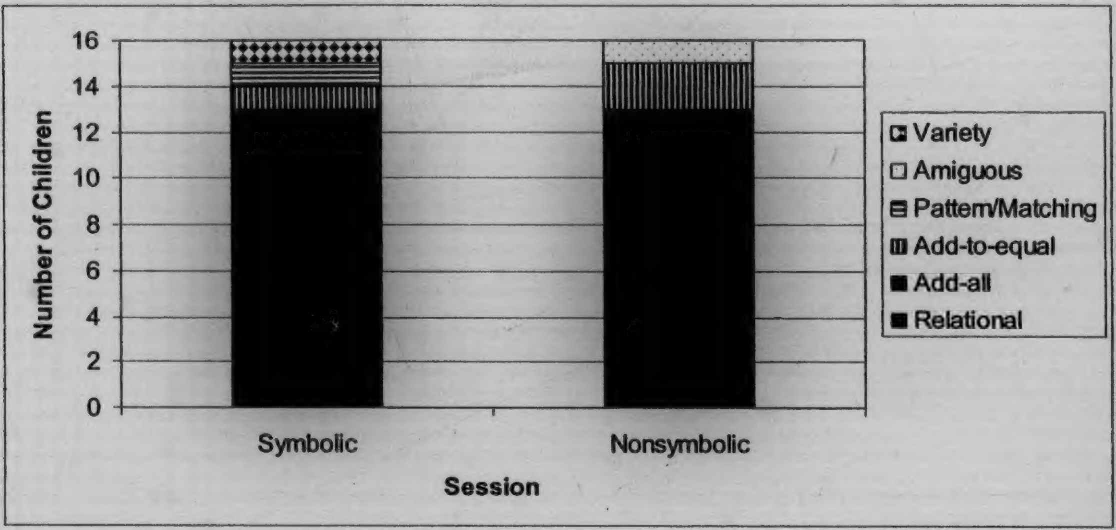


Figure 2-5. Number of the 16 Grade 2 children in the nonsymbolic/symbolic group who used specific strategies on both symbolic and nonsymbolic sessions.



Verbal Interpretations of the Equal Sign

None of the children in either group gave a relational verbal interpretation of the equal sign in the symbolic condition. Contrary to my prediction that children's verbal interpretations of the equal sign may differ, children who solved nonsymbolic problems first proceeded to describe the equal sign non-relationally in the arithmetic context a week later. Fifteen of the children in the symbolic/nonsymbolic group gave operator interpretations and one gave an ambiguous interpretation. Thirteen children from the nonsymbolic/symbolic group gave operator interpretations, two gave ambiguous interpretations, and one child said that she did not know what the symbol meant. This particular child, not surprisingly, did not solve any of the equivalence problems correctly in either condition, despite solving 92% of the arithmetic problems correctly. Thus children's performance on equivalence problems can be high even though they verbally describe the equal sign in operator terms in the context of an arithmetic equation. Perhaps a procedural improvement (i.e., using arithmetic to make both sides of an expression equivalent) precedes an ability to verbally define the symbol, or the procedure of solving relationally is more directly modified by exposure to nonsymbolic problems than verbally defining the equal sign.

Age-Related Changes in Performance

One notable result from previous research in this area (Sherman & Bisanz, 2004) was that Grade 2 children solving symbolic equivalence problems were no more accurate than kindergarten children solving the same problems nonsymbolically. The lack of an accuracy difference is significant developmentally, particularly when noted that Grade 2 children have had approximately two years of mathematical instruction and, arguably, plenty of exposure to symbolic arithmetic, compared to kindergarten children. It appears as though the presentation format (symbols versus objects) were enough to erase, in terms of accuracy, two years of experience with complex, symbolic and nonsymbolic, arithmetic (Sherman & Bisanz, 2004). I used similar analyses using the present data to determine whether that surprising result was replicated.

Seven of the equivalence problems that the kindergarten children solved in Study 1 were numerically identical to those presented to the Grade 2 children in Study 2, and the instructions given to the two age groups did not differ notably. I wanted to know whether accuracy on the equivalence problems was similar for kindergarten and Grade 2 children, particularly those Grade 2 children who solved symbolic problems first. Specifically, I wanted to compared Grade 2 children's performance on symbolic equivalence problems to kindergarten children's performance on the same problems presented nonsymbolically, despite the fact that the groups difference both by presentation format (symbolic vs. nonsymbolic) and approximately two years of age. A one-way ANOVA on the number of problems solved correctly for children in the three groups: Kindergarten ($M = 35.0\%$, $SD = 38.7\%$), Grade 2 symbolic/nonsymbolic ($M = 34.9\%$, $SD = 43.3\%$), and Grade 2 nonsymbolic/symbolic ($M = 67.0\%$, $SD = 46.9\%$) was conducted. Children's accuracy differed across the three groups, $F(2, 60) = 3.46$, $p < .05$. Contrasts were used to explore differences among the three group means. The performance of kindergarten students did not differ from that of the Grade 2 students in the symbolic/nonsymbolic group, replicating previous results (Sherman & Bisanz, 2005), and the mean for these two groups combined was lower than that for Grade 2 students in the nonsymbolic/symbolic group ($p < .05$). Remarkably, the presence of symbols was enough to make Grade 2 children's accuracy on symbolic problems no different from that of kindergarten children who solved nonsymbolic problems.

Discussion

In this study I confirmed what was only suspicion based on previous research (Sherman & Bisanz, 2004): Experience with nonsymbolic equivalence problems can actually lead to improvements on subsequent symbolic problems. This result is particularly surprising considering the fact that sessions were approximately one week apart and the experimenters made no direct references linking the two sessions together. Also exciting about this result is the potential for application to the classroom. Specifically, experience with the complex equivalence problems without the potentially misleading cues from conventional symbols (due to a misunderstanding of the equal sign) may help children map their successful approaches to symbolic problems. Having children learn to reason about equivalence problems in a nonsymbolic format would be a relatively easy, inexpensive intervention that could have a positive impact on children's ability to solve similar problems presented symbolically. Examining potential boundary conditions surrounding the effectiveness of this intervention would be an important step before suggesting the wide-spread use of manipulatives in the classroom as a method of presenting problems nonsymbolically and teaching children to think about equations relationally, however. There may have been other unique factors, in addition to the nonsymbolic presentation format, that helped children map their relational strategy to symbolic problems, such as the presence of multiple-choice options. Several such factors are explored in more detail in Chapter IV.

The within-subject design was instrumental for examining order effects, and detecting the impressive benefit of having children solve nonsymbolic problems prior to symbolic problems. Children in the nonsymbolic/symbolic group may have remembered the approach they used to solve the nonsymbolic problems and transferred the procedures to the same symbolic problems presented a week later. Perhaps children had no idea how to solve the symbolic equivalence problems (which is supported by the literature), so they used whatever they could construct from their strategy arsenal. For children in the nonsymbolic/symbolic group, their arsenal would consist of the strategies employed for problems they had seen a week before that seemed similar in some important and salient aspects, such as having quantities on both sides that needed to be considered when solving. Rittle-Johnson (2006) also found that, after a quick intervention, some children showed a significant improvement on equivalence problems, yet did not show similar improvements on tests of conceptual understanding of the equal sign. She too hypothesized that children's procedural approaches, and their self-explanations of those procedures, may have served as mechanisms of change even though children did not explicitly think about the conceptual rationale underlying the problems.

Although children in the nonsymbolic/symbolic group showed marked improvement on symbolic problems in their second session, children who began with symbolic problems had extremely low accuracy. The low performance of children in this group is consistent with previous research and helps support the argument that elementary school children are in need of specific instruction, intervention, and/or experience to help them not only solve equivalence problems successfully, but more importantly understand the equal sign in all, especially arithmetic, contexts. Use of the multiple-choice format was helpful in making judgments about the relative success, or in this case failure, on the equivalence problems. I recommend the use of such procedures for assessing children's performance on equivalence problems for this very reason. Furthermore, I was able to compare Grade 2 and kindergarten children's accuracy on a common subset of equivalence problems. Remarkably, I found no difference between Grade 2 children's performance on symbolic problems and kindergarten children's two age groups, although this result must be qualified by pointing out that the two groups were solving the same problems in two different conditions (nonsymbolic and symbolic).

Contrary to my initial hypothesis, children who solved nonsymbolic problems first did not give more relational verbal interpretations of the equal sign than children who solved symbolic problems first. In fact, no children from either group described the equal sign relationally. This outcome is surprising considering that children in the nonsymbolic/symbolic group performed significantly higher on symbolic problems than children from the symbolic/nonsymbolic group. Thus, children's performance on equivalence problems can be high even though they use operator terms to describe the equal sign in the context of an arithmetic equation. If the ability to give a verbal definition constitutes conceptual understanding, then perhaps the fact that children were able to *use* the equal sign in a relational manner yet not *explicitly define* the symbol relationally is another example of children changing procedurally before they change conceptually (Baroody, 2003; Rittle-Johnson, 2006; Rittle-Johnson & Siegler, 1998). However, with the present data I cannot speak to the order or type of understanding. Rather, children's disconnect between solving equivalence problems and verbally defining the equal sign may be another example of what Lawler (1981) referred to a *microworlds* of understanding. In his case study Lawler noted that his young daughter's accuracy on addition problems was significantly lower compared to when the same problems were presented with coins. That is, the child's experience with money made this "money world" of addition more available for solving arithmetic problems than experience with standard, paper-and-pencil addition problems. In the present study, children's *solving,* or procedural, world might be disconnected from the *defining* world, such that children were able to construct strategies for solving equivalence problems yet did not transfer their knowledge of equivalence to the task of defining the equal sign.

General Discussion

Together, the two studies expand upon previous research in at least two important ways. First, the results of Study 1 replicate and extend work done by Sherman and Bisanz (2004) demonstrating that some children enter school with an ability to use arithmetic in complex situations. Specifically, the overall average was higher than would be expected by random guessing and more than a quarter of the kindergarten children solved 50% or more of the nonsymbolic arithmetic equivalence problems correctly. This finding is consistent with the previous discovery that approximately a quarter (24%) of kindergarten children could convincingly solve equivalence problems in a relational manner (Sherman & Bisanz, 2004), but in this study the multiple-choice paradigm was useful for considering levels of success. Perhaps even more interesting than overall accuracy levels was the finding that kindergarten children tended to fit into three categories defined both by accuracy *and* agreement between solution choices and verbal justifications. Children who solved more than half of the equivalence problems correctly had high levels of agreement between solutions and justifications, most of which were relational. Children who solved all ten equivalence problems incorrectly were also easy to characterize, in this case as using operator strategies, as they too had a high degree of matching between solutions and justifications. Children who solved between one and four problems correctly, however, were more difficult to characterize as there was a high level of mismatches between solutions and justifications. Mismatches likely mean that children were either guessing or trying to maximize their likelihood of success by giving justifications that differ from their solution choices. This study provides the first conclusive evidence that children can invent both operator and relational strategies to solve equivalence problems before formal mathematical instruction. The same type of analysis could not be done in Study 2 because Grade 2 children's accuracy was not as variable and there were fewer students to categorize.

Study 2 provides the first evidence demonstrating that experience with equivalence problems in the nonsymbolic condition may facilitate performance on subsequent symbolic problems. The fact that accuracy on symbolic problems was so profoundly influenced has practical applications. Children may benefit in the classroom from practice with arithmetic equivalence problems presented without conventional symbols prior to solving problems presented symbolically. The nonsymbolic problems may help children map what they understand about arithmetic (which many of them may have had since at least kindergarten) to complex problems presented with symbols. Perhaps some qualification is necessary, however, because children in neither group verbally described the equal sign relationally. Children's performance on equivalence problems may be more rapidly amenable to change than their definitions of the equal sign in arithmetic contexts. Additional studies examining children's changing views of the equal sign before, during, and after exposure to equivalence problems in nonsymbolic (or other) contexts would help clarify potential relations between exposure to particular conditions, accuracy on equivalence problems, and verbal interpretations of the equal sign.

Results from the two studies were useful for considering potential interventions and instructional techniques, but they were also important for highlighting children's ability to invent procedures for solving complex mathematical problems and their propensity to map understanding onto symbols. Without feedback or prior formal instruction, kindergarten children often invented deliberate strategies to solve the equivalence problems. Some children invented a relational strategy, as indicated by solution choices and justification, leading to higher accuracy than would be expected by guessing. Others invented one or more operator strategies, such as adding all terms together, and their consistent, albeit incorrect, approach caused many children to solve all problems incorrectly. Operator approaches are common among children in elementary school and have been linked to experience with canonical arithmetic problems and a misunderstanding of the equal sign (McNeil & Alibali, 2005b; Seo & Ginsburg, 2003). Children may begin formal schooling with either a relation or operator strategy for solving equivalence problems. What is clear from other studies is that, when tested several years later, almost all children solve equivalence problems in an operator manner. How, when, and the types of experiences that change or reinforce children's strategies remains a direction for future research. Neither inventing a successful relational strategy nor inappropriately applying an operator strategy can be attributed entirely to formal instructional experiences or techniques. Study 2 provides clear evidence that children are capable of changing their strategies dramatically after relatively little exposure to problems in a nonsymbolic condition. It appears that children can create relational and operator strategies at a younger age than previously credited (Study 1) and they can change their strategies significantly after minimal practice on nonsymbolic problems, mapping successful procedures to problems that are typically failed in the symbolic context (Study 2).

How long Grade 2 children in the nonsymbolic/symbolic group maintained their improved performance on symbolic problems was not assessed in the present study; assessing the maintenance of these improvements is important for understanding the extent to which experience with nonsymbolic problems is beneficial. Similarly, further research is needed to help illuminate how children's accuracy on symbolic problems can be so drastically influenced by practice with nonsymbolic problems, yet their operator interpretations of the equal sign remain unchanged. Some of these issues are addressed in the following chapters. Practice with nonsymbolic problems may help children map relational strategies to symbolic problems as a whole, yet it may not be enough to override a prevalent, deep, misunderstanding of the symbol of equivalence. Clearly children's misunderstanding of the equal sign is resistant to change, at least with respect to indirect interventions such as solving problems with objects rather than conventional symbols.

References

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127-145.

Baroody, A. J. (2003). The developmental of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody and A Dowker (Eds.), *The developmental of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 1-33). Mahwah, MJ: Erlbaum.

Baroody, A. J., & Ginsburg, H.P. (1983). The effects of instruction on children's understanding of the "equals" sign. *The Elementary School Journal, 84*, 199-212.

Bisanz, J., Sherman, J. L., Rasmussen, C., & Ho, E. (2005). Development of arithmetic skills and knowledge in preschool children. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition (pp. 143-162).* New York: Psychology Press.

Davis, H., & Carr, M. (2002). Gender differences in mathematics strategy use: The influence of temperament. *Learning and Individual Differences, 13*, 83-95.

Felson, R. B., & Trudeau, L. (1991). Gender differences in mathematics performance. *Social Psychology Quarterly, 54*, 113-126.

Gelman, R. (1972). Logical capacity of very young children: Number invariance rules. *Child Development, 43*, 75-90.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teachers' hands tell the student's mid about math. Journal *of Educational Psychology, 91*, 720-730.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139-155.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L., Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly, 14*, 299-324.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*, 317-326.

Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology, 54,* 105-115.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equality and variable. *Zentralblatt für Didaktik der Mathematik – International Reviews on Mathematical Education, 37*, 1-9.

Lawler, R. W. (1981). The progressive construction of mind. *Cognitive Science, 5*, 1-30.

MacGregor, M., & Price, E. (1999). An exploration of aspects of language proficiency and algebra learning. *Journal for Research in Mathematics Education, 30*, 449-468.

McNeil, N. M., & Alibali, M. W. (2005a). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development, 6*, 285-306.

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466.

McNeil, N. M., & Alibali, M. W. (2002). A well-established schema can interfere with learning: The case of children's typical addition schema. In C. D. Schunn & W. Gray (Eds.), *Proceedings of the Twenty-fourth Annual Conference of the Cognitive Science Society* (pp. 661-666). Mahwah, NJ: Lawrence Erlbaum Associates.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology, 92*, 734-744.

Mix, K. S. (1999). Preschoolers' recognition of numerical equivalence: Sequential sets. *Journal of Experimental Child Psychology, 74*, 309-332.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6*, 449-468.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Rasmussen, C., & Bisanz, J. (2003). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91*, 137-157.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77*, 1-15.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*, 175-189.

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills: Studies in developmental psychology* (pp. 75-110). Hove, UK: Psychology Press.

Seo, K., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence…": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Sherman, J., & Bisanz, J. (2004, May). *Symbolic interpretation: Examining children's view of the equal sign*. Poster presented at the 16th Annual American Psychological Society, Chicago, Illinois.

Shrager J., & Siegler, R.S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*, 405-410.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Lawrence Erlbaum Associates.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skill* (pp. 229-294). Hillsdale, NJ: Lawrence Erlbaum Associates.

Western and Northern Canadian Protocol for Collaboration in Education (2006, May). *The common curriculum framework for K-9 mathematics*. Retrieved February 5, 2007, from http://www.wncp.ca.

CHAPTER III

JUST WHERE DO THEY GO WRONG? EXPLORING YOUNG CHILDREN'S MISUNDERSTANDING OF THE EQUAL SIGN BY MANIPULATING THE DEGREE OF SYMBOLIZATION IN EQUIVALENCE PROBLEMS

Basic literacy and mathematical skills depend in part on children's ability to interpret and use symbols such as letters, numbers, and operator symbols (e.g., + and -). Mathematics has proven to be an excellent area in which to examine children's use of strategies and understanding of symbols when solving problems (Miller, 2000). Examining solution strategies, and the reasons children give for solving problems in specific ways, can inform us about the concepts they understand and factors that may constrain their understanding. One such factor is the format in which mathematical problems are presented. Because arithmetic is typically expressed using symbols, examining how children understand or misunderstand these symbols (i.e., their mathematical literacy skills) has important implications for both educational instruction and for assessing academic achievement.

Development of Symbol Use

Before children can use symbol systems such as letters or numbers, they must realize that such symbols stand for something and that this relation is constant. Bialystok and Martin (2003) conducted three studies with 4-year-old pre-readers using various forms of the *moving word task* to determine the extent to which children realized that words (symbols expressing language) had permanent relations to corresponding pictures. In one task an "accident" occurred such that the word card was moved from below the congruent picture to a position below an incongruous picture. When asked what the word said, children overwhelmingly indicated that the word had changed once it was placed under the incongruent picture. Bialystok and Martin concluded that "until children understand how these notational systems carry the meaning they encode, their knowledge of them is representational, but it is not symbolic" (p. 227). That is, children were representing the words in terms of the photographs they were next to, rather than viewing a word as a symbol whose relation with its referent remains unchanged regardless of context.

Hughes (1986) also examined young children's ability to use written symbolism, particularly to determine whether stages of symbolism exist, and studied mathematical symbols rather than letters. Although written symbols, such as numerals and letters, are typically taught in school, Hughes' work suggests that children begin to appreciate the usefulness of representing quantity prior to formal instruction. By asking three- to seven-year-olds to "put something on paper" to represent 1, 2, 3, 5, or 6 bricks, Hughes found evidence for four categories of written representation. The idiosyncratic category was characterized by markings that did not appear to reflect any numerical information (e.g., scribbling), whereas the pictographic category consisted of markings that represented both the appearance and numerosity of the blocks (e.g., pictures of *x* number of blocks). The iconic category included markings, such as tally marks, that demonstrated one-to-one correspondence with the number of blocks but lacked visual cues such as the shape of the blocks. In the symbolic category, children used conventional symbols (e.g., "5") to represent the number of blocks. In general, the younger children gave idiosyncratic or pictographic markings, whereas the older children used iconic or symbolic depictions of numerosity (also see Munn, 1998).

As evident above, children can use symbols to represent numerosity prior to formal instruction. Despite general success in representing static sets of bricks, Hughes (1986) found that six- to ten-year-olds have much less success representing operations, such as addition and subtraction. In fact, not one of the children he asked could adequately represent arithmetic such as 2 + 2, and none of the children used the conventional "+" or "-". Hughes concluded that even preschool children can represent small numbers (often with one-to-one correspondence), yet by age 10 most children still have trouble representing operations. He noted a reluctance to use conventional signs of arithmetic and indeed, he stated that "the impression which [he] got from several of the children was that the signs '+', '-', and '=' had little immediate relevance to the world of ordinary objects, but inhabited the self-contained world of 'sums'" (p. 103).

Understanding the Symbol of Equivalence

Although children typically learn to master written numerals and even most operator symbols relatively quickly after receiving formal instruction, some symbols appear to take longer to appreciate. One example in particular is the symbol for equivalence, or the equal sign, which children notoriously misinterpret throughout elementary school (Alibali, 1999; Baroody & Ginsburg, 1983; Behr, Erlwanger, & Nichols,1980; Kieran, 1981; Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; McNeil & Alibali, 2000, 2002, 2004, 2005a, 2005b; Perry, 1991; Perry, Church, & Goldin-Meadow, 1988; Renwick, 1932; Rittle-Johnson & Alibali, 1999; Sáenz-Ludlow & Walgamuth, 1998; Seo & Ginsburg, 2003), and possibly into early adulthood (Kieran, 1981). This misinterpretation is generally detected by children's failure to solve equivalence problems such as 4 + 7 + 3 = 4 + __. Even Grade 4 or 5 students overwhelmingly respond by adding all numbers (i.e., 18) or by adding those to the left of the equal sign (i.e., 14) and putting those numbers on the line, rather than making both sides of the expression quantitatively identical. Children often view the equal sign as an indicator to "do something" (an *operator* view) rather than a symbol of equivalence (a *relational* view). Unfortunately, misunderstanding the equal sign can cause serious difficulties such as rejecting noncanonical problems and thinking they are written "backwards" (e.g., __= 2 + 3), as well as in higher mathematics such as algebra (McNeil & Alibali, 2005a).

There is some evidence that children understand the equal sign in contexts other than arithmetic equations (such as between a dime and two nickels), suggesting that their understanding is contextually bound (Seo & Ginsburg, 2003). When presented in arithmetic problems with conventional symbols, children misinterpret the symbol as being operator in nature (i.e., "do something"). When presented in non-arithmetic contexts, or contexts without conventional mathematical symbols, children interpret the equal sign in a relational manner (i.e., "same as", or "equal to"). Recent research suggests that it is the presence of conventional mathematical symbols and not the context of arithmetic problems per se that leads to poor performance (Sherman & Bisanz, 2004, Chapter II).

The effect of the important distinction between conventional symbols versus arithmetic context was tested by presenting arithmetic equivalence problems using objects rather than symbols (Sherman & Bisanz, 2004, Chapter II). In both studies, problems were presented using blocks in small bins to represent the addends with a cardboard divider (a blue "tent") to represent the equal sign (*nonsymbolic*, Appendix A).

The objects allowed students to view and solve the problems without the use of typical symbols such as "4" or "=." The structure of the equivalence problems took several formats, most notably *part-whole* (e.g., $a + b = c +$ __) and *combination* (e.g., $a + b + c = a +$ __) problem types. The two types of problems potentially required different processing constraints: Part-whole problems could only be solved using arithmetic, whereas combination problems could be solved by ignoring the two *a* terms and simply adding the *b* and *c* terms and putting the sum on the line (as discussed in Chapter II).

In the first study (Sherman & Bisanz, 2004), Grade 2 and kindergarten children were instructed to put blocks in the last (empty) bin to make both sides of the blue tent have the same total number of blocks. Grade 2 children were randomly assigned to solving equivalence problems in either the nonsymbolic condition or a typical, symbolic condition. In the second study (Chapter II, Study 2), students were asked to solve similar problems by selecting from several photographs of blocks and placing the chosen photograph in the last (empty) bin. The photographs corresponded to answers that would be generated by specific, well-documented, strategies for solving equivalence problems such as the sum of *all* addends (add-all), or the sum of the addends to the left of the equal sign (add-to-equal). The multiple-choice paradigm of the second study was useful for comparing performance to chance levels of accuracy to determine whether children in various groups (e.g., kindergarten versus Grade 2) were correct on a convincing number of problems to be categorized as successful. Grade 2 children in this study were randomly assigned to either solving nonsymbolic problems in the first session, followed by the same problems presented symbolically a week later (nonsymbolic/symbolic), or the opposite order (symbolic/nonsymbolic). In both studies, kindergarten children received only nonsymbolic problems.

In both studies (Sherman & Bisanz, 2004, Chapter II, Study 2), Grade 2 children had much higher accuracy on nonsymbolic problems than on the same problems presented with conventional symbols. Despite the complexity of the arithmetic involved, even some kindergarten children demonstrated high levels of success on the nonsymbolic problems. In fact, more than a quarter of kindergarten children (30%) solved at least half of the nonsymbolic problems correctly, and overall the children solved more problems correctly than would be expected if they were merely guessing (Chapter II, Study 1). Thus, many children appear to begin formal schooling already possessing the ability to use arithmetic to make two sides of an arithmetic problem equivalent. Perhaps the most interesting result from the research was the order effect found in the Study 2. Children in the nonsymbolic/symbolic group not only had extremely high performance on the nonsymbolic problems, they appeared to benefit from having experience with nonsymbolic problems first. Their performance on subsequent symbolic problems was much higher than for children who solved symbolic problems first.

The studies in Chapter II were instrumental in demonstrating that: (a) Grade 2 children do markedly better on equivalence problems when they are presented without mathematical symbols than with symbols, (b) when Grade 2 children experience equivalence problems in the nonsymbolic condition first, they do much better on subsequent symbolic problems than if they start with symbolic problems, (c) Grade 2 children appear to solve both problem types (part-whole and combination) with equal facility, and (d) kindergarten children solve equivalence problems with higher overall accuracy than would be expected with guessing, and in fact almost 20% of the children solved all 10 problems correctly, and (e) children in elementary school may have difficulty mapping what they know about using arithmetic on equivalence problems onto problems presented with symbols.

Schemas of Typical Arithmetic Patterns

What is it about symbolic equivalence problems that make them so difficult to solve? Although poor performance is often attributed to the equal sign in particular, another hypothesis is that it is the arithmetic-looking context more generally (i.e., when the entire problem is presented symbolically), rather than the equal sign specifically, that leads to low accuracy. McNeil and Alibali (2005b) discovered that children may be vulnerable to *change resistance* as a result of exposure to symbolic arithmetic problems primarily presented in a canonical format (i.e., $a + b =$ __). Practice with problems of this form only, whereby the answer immediately follows the equal sign, may enforce a belief that the equal sign means "the answer comes next" rather than that both sides are numerically identical. As Seo and Ginsburg (2003) discovered, teachers' explanations and typical mathematical textbooks largely consist of canonical problems and examples that reinforce the equal sign as signaling the end of an arithmetic expression rather than equivalence. Therefore, much like applying a *mental set*, children may approach arithmetic equivalence problems using their well-rehearsed strategy of adding all terms together and putting the sum on the blank line, rather than ensuring that both sides of the equal sign are quantitatively equivalent.

The structural appearance of typical, symbolic, equivalence problems may appear sufficiently similar to typical canonical problems for children to apply their mental set of adding from the left and putting the sum on the blank (rather than ensuring equality on both sides of the equal sign). Using this type of mental set is an efficient method for solving problems that fit a canonical format. The benefits disappear, however, when the format of the problem differs, as is the case with equivalence problems. According the change resistance account, the more experience children have with canonical problems, and the stronger children's adherence to an *operator* view of the equal sign, the more likely children are to fail symbolic equivalence problems. Children's failure on equivalence problems could be due to several factors. First, children may not have the arithmetic skills necessary to solve such complex problems. We discovered this was not the case, however, as even many kindergarten children could solve these challenging problems (Chapter II; Sherman & Bisanz, 2004). Second, it could be that, structurally, equivalence problems look sufficiently enough like canonical problems to activate existing mental sets. The form of the problem (i.e., a number of addends, separated by operator symbols such as "+", and a blank line at the end), or perhaps the mere presence of conventional mathematical symbols, may be what triggers a prepotent response for "doing something" to solve. In this case, children would apply a sort of addition schema, mindlessly solving from left to right (add-all), ignoring the position of the equal sign. The Grade 2 children in the nonsymbolic/symbolic condition (Chapter II, Study 2) showed that, at least under the right conditions, namely practice with nonsymbolic problems first, children are able to override the potential mental set and solve symbolic equivalence problems with relatively high levels of success.

In summary, children overwhelmingly fail equivalence problems in the presence of conventional symbols (except perhaps when they have had prior experience with nonsymbolic problems), whereas in the absence of all conventional symbols children overwhelmingly succeed on the very same problems. Left unaddressed is the extent to which the structural form of the equivalence problems, and presence of symbols, influence children's responses. How will children's view of the equal sign and/or their mental set for arithmetic problems influence how they reason about equivalence problems presented with just some conventional symbols? I wanted a condition in which equivalence problems may or may not look sufficiently like typical arithmetic problems to determine whether children apply their mental set (and solve problems incorrectly), override their mental set (and solve problems correctly), or even benefit from having to reason about the strategies and the equal sign in the novel condition (and solve problems correctly *as well as* show improvements on subsequent problems presented in the typical, symbolic, condition).

For this study, I created a novel condition in which equivalence problems were presented with only some of the mathematical symbols found in typical symbolic problems. This condition was designed to serve as an intermediate condition between the nonsymbolic and symbolic conditions used in Chapter II (Study 2). In this semi-symbolic condition, the equal sign was present, as were the addition signs, yet the addends were represented without Arabic symbols. If children have low accuracy on these problems, it may be because of their misunderstanding of the equal sign. That is, even though semi-symbolic problems may not look like typical arithmetic problems, the equal sign is still present and disruptive, and may therefore negatively influence accuracy. In this case, the equal sign alone may be enough to evoke change resistance. If children have high accuracy on these problems, then perhaps the equal sign alone is not enough to cause the difficulty. Rather, the problem may be due to the broader arithmetic context when all parts of the equation are represented using symbols.

In light of the order effect from Chapter II (Study 2), I used a similar within-subject design to determine whether performance on symbolic problems would differ for children who solved semi-symbolic problems first versus symbolic problems first. Similarly, the design allowed me to examine whether solving symbolic problems first would influence performance on semi-symbolic problems. Finally, strategy use, identified using as a match between solution choice and justification (e.g., an add-all solution and an add-all justification), was examined to determine whether children's approaches differed in relation to equivalence session order, and was compared to strategy use by children in Chapter II, Study 2. Matches provide convincing evidence that children deliberately chose the solutions and could explain why they did so.

Method

Participants and Design

Participants consisted of 32 Grade 2 students, including 18 males and 14 females, from three public schools. Ages ranged (in years; months) from 7;3 to 8;5, with a mean of 7;9. All children received two 20-minutes sessions, with approximately one week between sessions, both of which began with a variety of practice problems used to familiarize them with the materials and instructions. Each child was tested by the same experimenter for both sessions.

To assess whether performance differed across the two conditions, children received the same 16 problems in both the symbolic and semi-symbolic contexts. To determine whether children's performance was affected by order of experience within the two conditions, children were randomly assigned to one of two session orders: symbolic followed by semi-symbolic (symbolic/semi-symbolic), including 9 males and 7 females, or semi-symbolic followed by symbolic (semi-symbolic/symbolic), including 9 males and 7 females.

Materials and Procedures

Symbolic condition. Problems were presented using a black stand placed on top of a table or desk. White laminated cards containing the problems were flipped over one at a time. Prior to the presentation of each problem, solution choice cards were arranged along a horizontal line beside the problem card, fixed with Velcro (Appendix A). Each problem contained a blank line and Velcro strip to indicate where the solutions needed to be placed. Children solved by removing the solution card of their choice from the line and fixing it to the blank line on the problem card. The choice cards consisted of quantities (Arabic numerals) that corresponded to four specific ways of approaching the equivalence problems (add-all, add-to-equal, relational, and choosing the smallest quantity) (see explanation in Chapter II, Study 2).

At the beginning of both sessions, children received practice problems that included a counting string (1, 2, 3, 4, __, 6), one basic arithmetic problem (2 + 1 = __, with 2 + 2 = __ as a back-up problem), and a practice equivalence problem (1 + 1 = 1 + __). For all arithmetic and equivalence problems, children were instructed: "put a card on the line so that when you put together these (experimenter points to terms on the left side of the equal sign) you'll have the same number on this (points left) side of the equal sign as when you put together these (experimenter points to the addend and the blank line on the right side of the equal sign) on this (points right) side of the equal sign." After solving each practice problem, feedback was given and children were asked to justify their response by telling the experimenter why his or her answer was correct. Justifications were coded using the criteria described in Appendix B. If justifications were coded the same as the corresponding solution choice (e.g., an add-to-equal justification and the add-to-equal solution choice), the response was considered a match, with matches being strong evidence of strategy use.

Following the practice problems, the problem set (consisting of 16 equivalence problems, see Appendix C) was administered. Instructions were given for the first two problems, students were asked to justify their responses, and solutions, any observable solving behaviors, and verbal justifications were recorded. No feedback was given. In the problem set, the order in which the choice cards were presented was determined using a Latin Square for the first and second half (8 problems each) of the problem set. Order of choice cards was never the same for two questions in a row. Half the problems were part-whole types and half were combination types, and there were never more than two of the same problem types in a row.

Semi-symbolic condition. Semi-symbolic problems were presented similarly to symbolic problems. A black stand was used, and problems appeared on white laminated cards that were flipped over one at a time. As in the symbolic condition, children responded by choosing one of five choice cards (four values corresponding to specific strategies for solving and the question mark). The choice cards and all addends in the equivalence problems were represented using dots rather than Arabic numerals, however. For example, "2" was represented as "• •". All other conventional symbols, specifically "+" and "=" were used in typical fashion (Appendix A). As described above, children solved practice problems first (with feedback), followed by the problem set (16 equivalence problems).

For both the symbolic and semi-symbolic sessions, scores for the problem set were converted to percent correct because of experimenter error (as in Chapter II, Study 2). That is, eight children received only 15 equivalence problems due to a duplicate problem in the problem set. Once the duplicate was detected, a new problem was created for the remaining participants.

Arithmetic problems. Completion of the problem sets for both symbolic and semi-symbolic sessions was followed by six arithmetic problems (4 + 3 = __, 6 + 4 = __, 1 + 3 = __, 2 + 7 = __, 3 + 5 = __, and 2 + 3 = __), presented in the same condition as the problem set. Both symbolic and semi-symbolic problems were solved by children writing their solutions on the problem cards using a dry-erase pen. In the semi-symbolic condition, children were encouraged to respond using either dots (corresponding to how the addends were represented) or Arabic numerals. The problems were within the same magnitude range as the equivalence problems, and were designed to assess addition skill level so that failure on equivalence problems could not be blamed on poor arithmetic skills in general. No feedback was given.

Verbal interpretation of the equal sign. Following completion of the arithmetic problems, children were asked what the equal sign means. Specifically, the experimenter would cover both sides of the last arithmetic problem (2 + 3 = __) in both the symbolic and semi-symbolic sessions, leaving just the equal sign exposed, and ask “what is this?”, followed by “what does it mean?” Children’s explanations were coded (Appendix D), and no feedback was given. Children were prompted to give respond again if they stated that the equal sign means "equals."

Results and Discussion

Accuracy

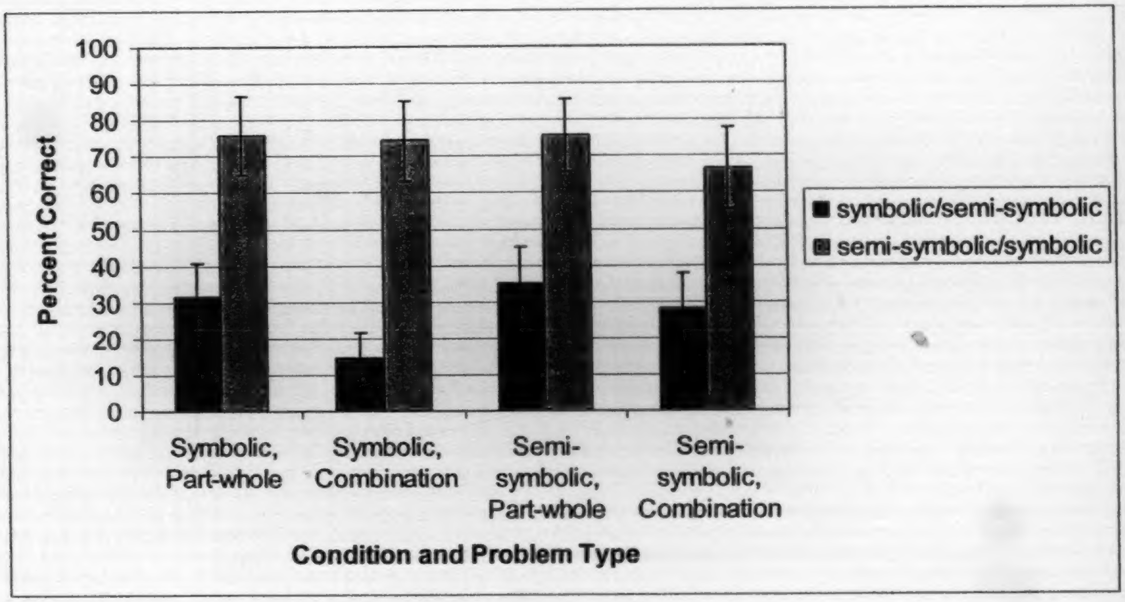
A 2(Gender) x 2(Session order: symbolic/semi-symbolic, semi-symbolic/symbolic), x 2(Condition: symbolic, semi-symbolic) x 2(Problem Type: part-whole, combination) ANOVA with repeated measures on the last two variables was conducted to explore effects and interactions related to gender, session order, and condition. Children who received semi-symbolic problems first (semi-symbolic/symbolic) did much better on average ($M = 74.6\%$) than children who received symbolic problems first (symbolic/semi-symbolic) ($M = 26.5\%$), $F(1, 28) = 13.73, p < .01$ (Figure 3-1).

Interestingly, accuracy did not differ between the symbolic and semi-symbolic conditions, $F(1, 28) < 1$. It was not the case that semi-symbolic problems were easier to solve than symbolic problems, unlike the pattern of results with nonsymbolic problems in Sherman and Bisanz (2004) and Chapter II. Rather, session order mediated success on problems in both conditions. Children who solved semi-symbolic problems first went on to do quite well on subsequent symbolic problems, such as the children in the nonsymbolic/symbolic group in Chapter II. In contrast, children who began with symbolic problems had low accuracy on the semi-symbolic problems presented a week later. As with the nonsymbolic condition in Chapter II (Study 2), it appears as though solving problems in the semi-symbolic condition facilitated performance on subsequent symbolic problems, even though the sessions were approximately a week apart and experimenters did not deliberately or explicitly draw children's attention to the similarity between the two "tasks" across sessions.

In contrast, children who begin with symbolic problems continue to perform poorly on semi-symbolic problems presented a week later. Perhaps the initial exposure with symbolic problems activates an addition schema that continues unchallenged on subsequent semi-symbolic problems. Unlike the second study in Chapter II, where the alternate condition (i.e., nonsymbolic problems) looked very different from the symbolic condition, semi-symbolic problems may look sufficiently like symbolic problems to the extent that children apply their mental set to solve them once the mental set has been activated.

In addition to the session order effect, children had higher accuracy on part-whole problems (55.0%) than on combination problems (46.1%), $F(1, 28) = 6.29, p < .05$, similar to the pattern found with kindergarten children but not the Grade 2 children in Chapter II. This effect may be due to the exceptionally poor performance on symbolic combination problems by children in the symbolic/semi-symbolic condition. Reasons for this result are not clear, and it was not found in Study 2 (Chapter II). In any event, this interaction does not compromise the most important finding that session order, but not condition, influences accuracy. No other comparisons or interactions were significant. Scores on the six arithmetic problems were extremely high ($M = 5.91$, $SD = .30$ in the symbolic condition; $M = 5.97$, $SD = .18$ in the semi-symbolic condition) and will not be discussed further.

Figure 3-1. Mean accuracy by Grade 2 children (symbolic/semi-symbolic and semi-symbolic/symbolic session order groups) on part-whole and combination equivalence problems presented in both symbolic and semi-symbolic conditions, with standard error bars.



The multiple-choice method was used to compare accuracy rates to chance levels and to examine children's solving strategies for consistent patterns (i.e., did children choose the add-all solutions on most trials or were they guessing randomly?). With respect to the first benefit, the probability of randomly choosing any one response option was considered to be .2 because of the five solution options given on every trial (add-all, add-to-equal, relational, small number, and other). For children in the symbolic/semi-symbolic group, performance was low on symbolic problems (*M* = 22.5%, *SD* = 29.9) and on semi-symbolic problems (*M* = 31.4%, *SD* = 36.2). In both cases the 95% confidence interval included 20%, indicating that the obtained means were not reliably different than a level that would be obtained by guessing. Thus, children in the symbolic/semi-symbolic group had accuracy that was no higher than could be expected if children guessed on each problem. In contrast, children from the semi-symbolic/symbolic group had mean accuracy levels exceeding 70% for both conditions, with 95% confidence intervals producing ranges far higher than could be expected by guessing.

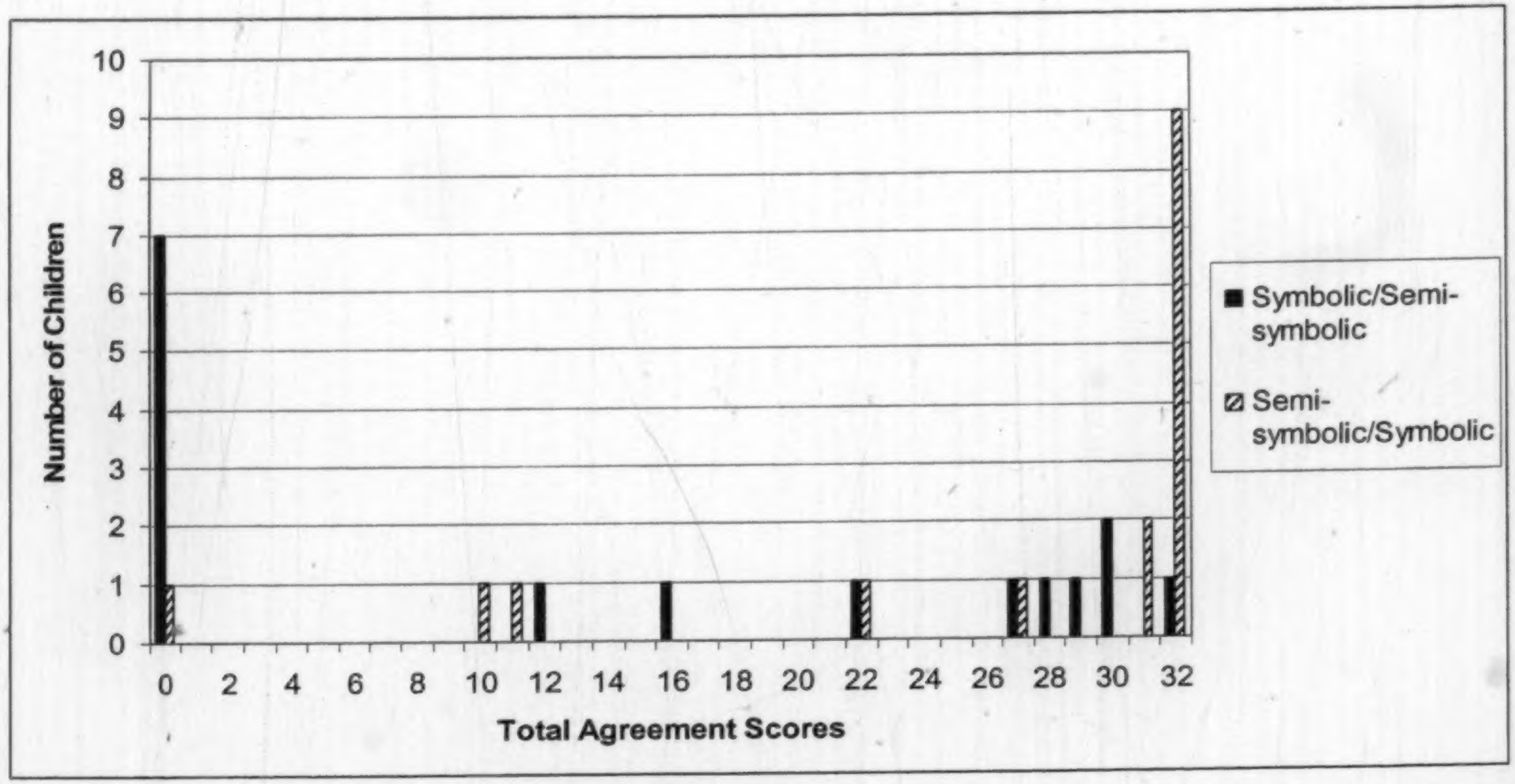
Strategy Use

Overall agreement between solutions and justifications was relatively low (63.3%), prompting detailed analyses of solutions, justifications, and situations in which solutions and justifications did (or did not) match, rather than using solution options alone. When the overall agreement between solutions and justifications is low, there are limitations to examining only solutions or justifications in isolation as evidence of strategy use, as described in Chapter II (Study 1). Mismatches between children's coded justifications and their solution options may be evidence of guessing or of trying multiple approaches, perhaps within a single problem. By selecting a solution option that represents one approach, while justifying a different approach, children may have been attempting to increase their chances of looking or sounding correct. Therefore I examined just the cases where children's actions and statements did match. Matches were considered to be convincing evidence of strategy use.

Agreement scores varied widely, with 10 children having perfect agreement across all 32 problems (nine children from the semi-symbolic/symbolic group, one child from the symbolic/semi-symbolic group), and eight children with an agreement score of zero (seven from the symbolic/semi-symbolic group, one from the semi-symbolic/symbolic group) (Figure 3-2). Thus, the lowest end of the agreement spectrum consisted largely of children who solved symbolic problems first, whereas the highest end of the agreement consisted mainly of children who solved semi-symbolic children first. Therefore, children who received semi-symbolic problems in the first session were not only more likely to have high accuracy on both sessions compared to children who began with symbolic problems, but they were also more likely to support their solution options with matching justifications (i.e., mainly relational).

Due to the variability of accuracy and agreement scores related to the order in which children received the sessions, strategy use was examined separately for the symbolic/semi-symbolic and semi-symbolic/symbolic groups. I also compared strategy use on symbolic problems by children in the symbolic/semi-symbolic group to children symbolic/nonsymbolic children from Chapter II, Study 2, using two methods. First, I examined matches between solutions and justifications as evidence of strategy use for children in both groups, as described above. Second, because children's strategy use in the symbolic/nonsymbolic group (Chapter II, Study 2) was classified using solution choices alone (i.e., at least 13 of the same solution type across the 16 problems per session), solution choices alone were considered using the 13-of-16 criterion as in the previous chapter. If children responded in the same fashion on at least 13 of the 16 problems (approximately 80%), they were classified as solving in that manner. For example, if a child chose the add-all solution choice on 13 or more problems, he or she was classified as using an add-all strategy. The binomial probability of randomly selecting the same type of response, with an α of .05 and the probability of choosing any one response being .2, is less than .0001. Because children in both the symbolic/semi-symbolic and symbolic/nonsymbolic groups received symbolic problems first and mean accuracy was low in both groups, I expected that differences in strategy use would be minimal.

Figure 3-2. Distribution of total agreement scores (with a maximum score of 32 matches between solutions and justifications) for the 32 Grade 2 children in the symbolic/semi-symbolic and semi-symbolic/symbolic conditions.



The 16 children in the symbolic/semi-symbolic group had 224 matches between solution options and justifications across the 32 equivalence problems for an overall agreement of 44.1%. Approximately half (47.3%) of all matches involved relational strategies, and the rest (52.7%) were operator strategies. Seven of these 16 children had no matches on any of the problems. Children demonstrated an increase in the frequency of relational strategies from the first (symbolic) to the second (semi-symbolic) session, as evident in Figure 3-3. In contrast, the 16 children in the semi-symbolic/symbolic group had 422 matches between solutions and justifications across the 32 equivalence problems, for an overall agreement of 82.4% across all problems. Most of the matches represented the relational strategy (87.7%), with very few operator strategies (12.3%) (Figure 3-4). Thus, children from both session order groups demonstrated add-all, add-to-equal, and relational strategies, but children who solved symbolic problems first used far fewer relational strategies, and had fewer matches between solutions and justifications, than children who began with semi-symbolic problems.

Figure 3-3. Percentage of strategy types, as defined by matches between solutions and justifications, on symbolic (top panel) and semi-symbolic (bottom panel) sessions by children in the symbolic/semi-symbolic group.

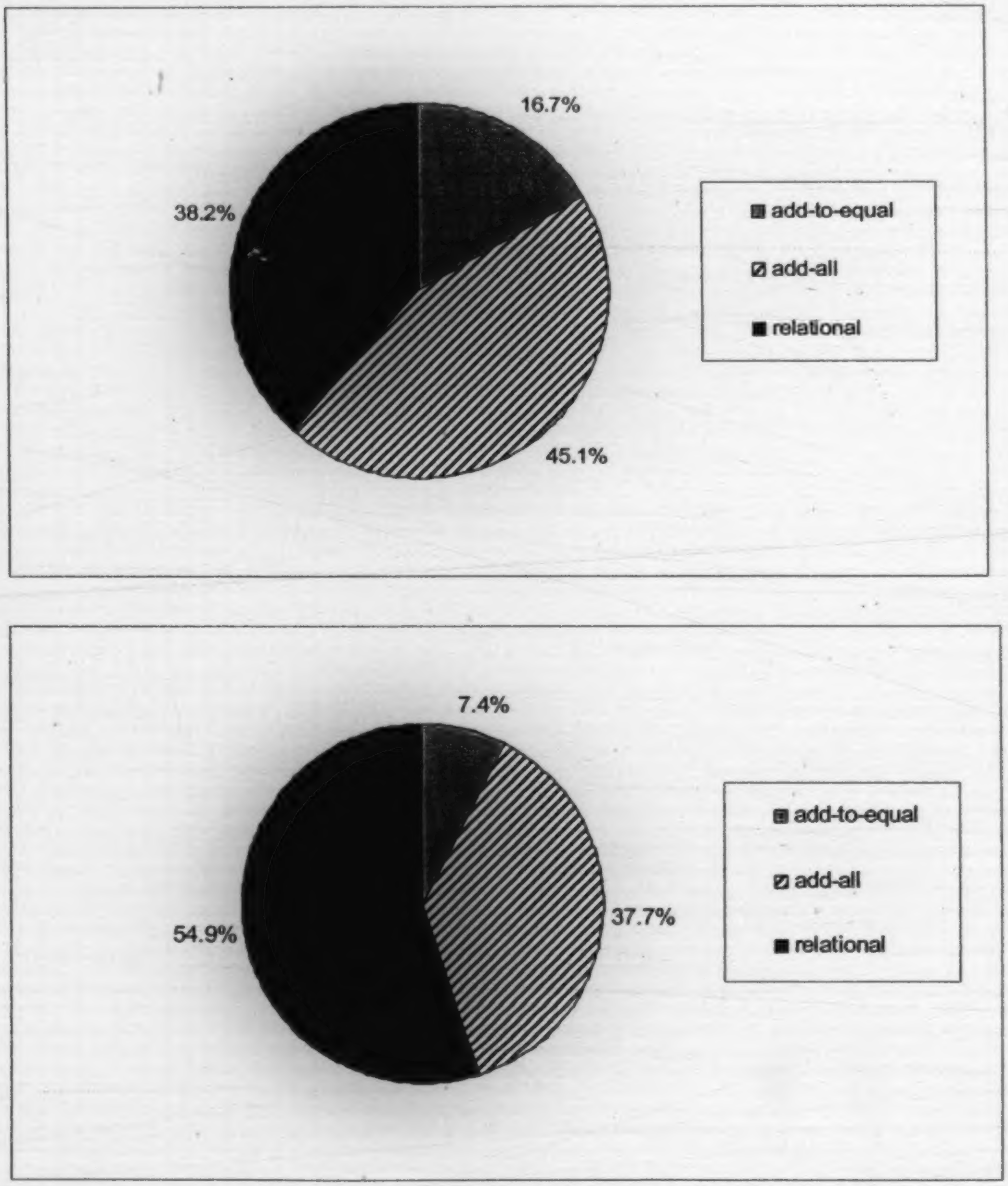
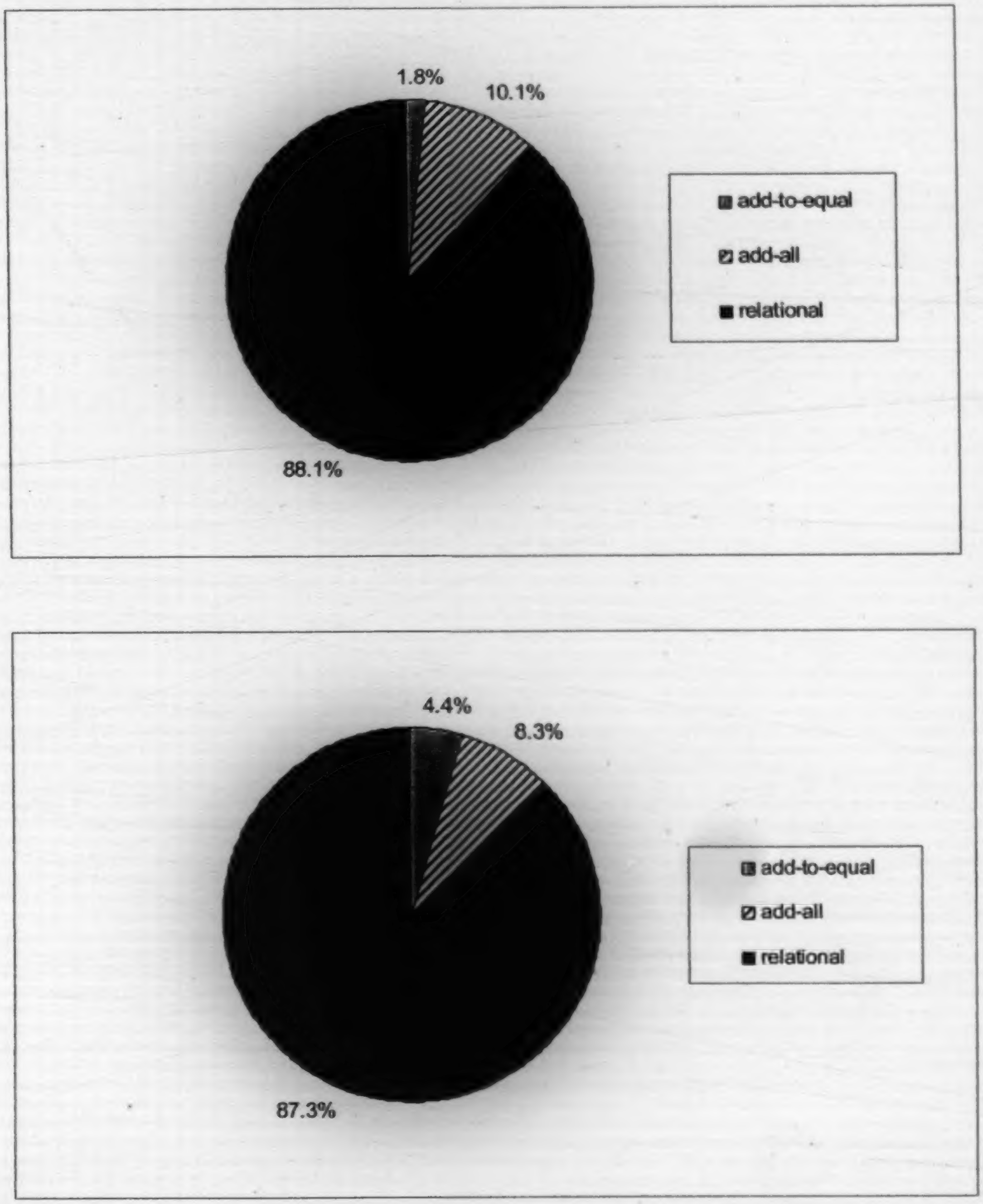


Figure 3-4. Percentage of strategy types, as defined by matches between solutions and justifications, on symbolic (top panel) and semi-symbolic (bottom panel) sessions by children in the semi-symbolic/symbolic group.



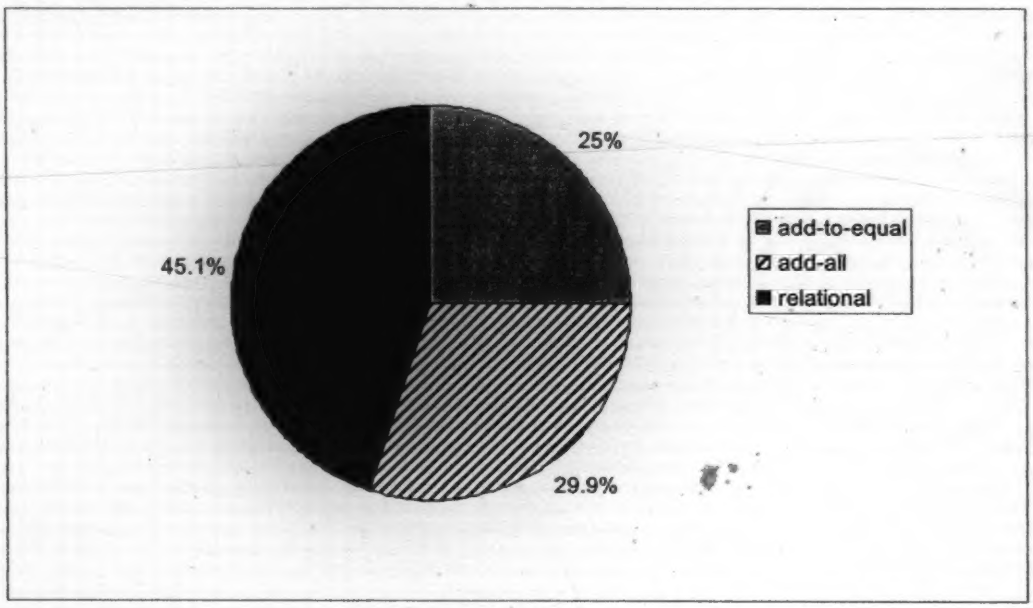
Children from the symbolic/semi-symbolic group and symbolic/nonsymbolic group (Chapter II, Study 2) had two trends in common. First, accuracy (and hence relational strategies) among children from these two groups was substantially lower than accuracy among children from the two opposite-order groups, at least on symbolic problems. Second, children had fewer matches between solutions and justifications compared to children from the opposite-order groups. Children in the symbolic/semi-symbolic order had less than 50% agreement overall, compared to 82.4% agreement for children in the semi-symbolic/symbolic order. Similarly, children in the symbolic/nonsymbolic group had approximately 74% agreement overall compared to 85.6% by children in the nonsymbolic/symbolic order. When considering agreement scores for the symbolic session only, the average agreement percentages for the two groups are more similar: 40.6% for the symbolic/semi-symbolic group and 63.3% for the symbolic/nonsymbolic group.

The number of relational and operator matches on symbolic problems were compared between children in the symbolic/semi-symbolic and symbolic/nonsymbolic groups (Figure 3-5). Relational strategies made up less than 50% of the matched cases by both groups of children. Second, children were credited as consistently using a specific strategy, such as relational or add-all strategies, if they chose the same type of solution on at least 13 of the 16 symbolic problems, regardless of matching or mismatching justifications. This method of classifying strategy users revealed the wide variability with respect to the solution choices children consistently used when solving symbolic problems (Figure 3-6). Less than half of the children in either group could be classified as relational strategy users. As hypothesized, children who began with symbolic problems approached the problems similarly, choosing a variety of solution options, agreeing on solutions and justifications for operator strategies more than the relation strategy, and having low accuracy relative to children from the opposite-order groups.

Verbal Interpretations of the Equal Sign

None of the children in either group gave a relational verbal interpretation of the equal sign when asked to define what "=" meant. On both the symbolic and semi-symbolic problems, 88% of the verbal responses given were operator, with the remaining responses being classified as "I don't know," much as in Chapter II, Study 2. In one case of experimenter error, the question was not administered to the child. Children's verbal interpretations of the equal sign do not appear to be as modifiable by exposure to problems in contexts other than symbolic as their performance on equivalence problems. The discrepancy between accuracy on equivalence problems versus verbal definitions does not mean that performance on equivalence problems and views of the equal sign are independent, however. In fact, there is research to show they are very much related (Knuth et al., 2005). What I have shown, however, is that practice with equivalence problems with no or just some mathematical symbols first can help accuracy on symbolic equivalence problems at a later date, without changing children's definitions of the equal sign. Accuracy (procedures for solving equivalence problems) and verbal interpretations (ability to explain what the symbol for equivalence means) may be sensitive to change at different rates, and may be impacted by different types of experience. Intervention in the opposite order, however, may lead to higher rates of giving relational definitions of the equal sign. That is, deliberate attempts to teach students relational views of the equal

Figure 3-5. Percentage of strategy types on symbolic problems, as defined by matches between solutions and justifications, for children in the symbolic/nonsymbolic (top panel, Chapter II, Study 2) and symbolic/semi-symbolic (bottom panel) groups.



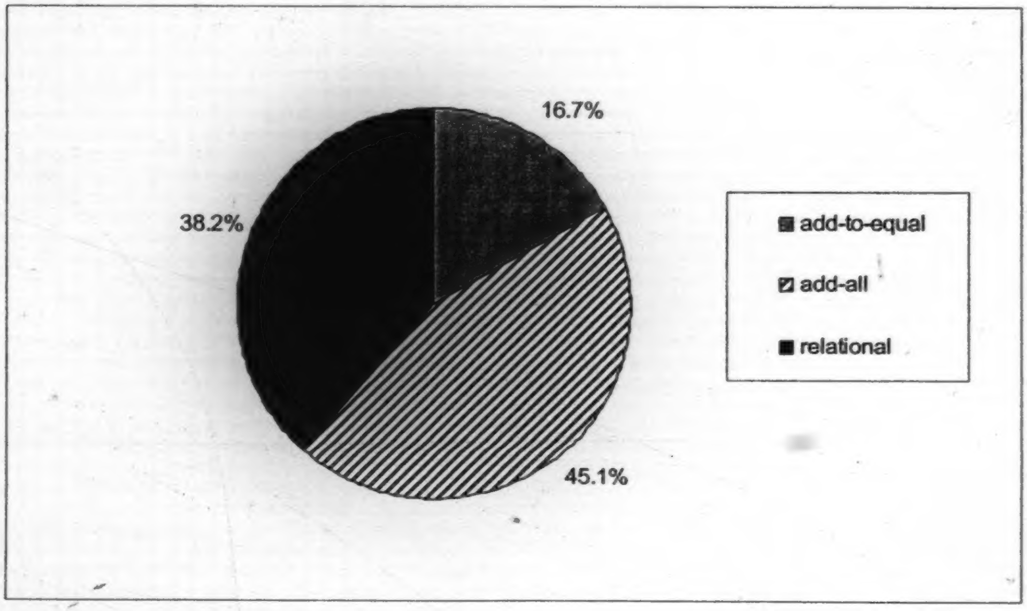
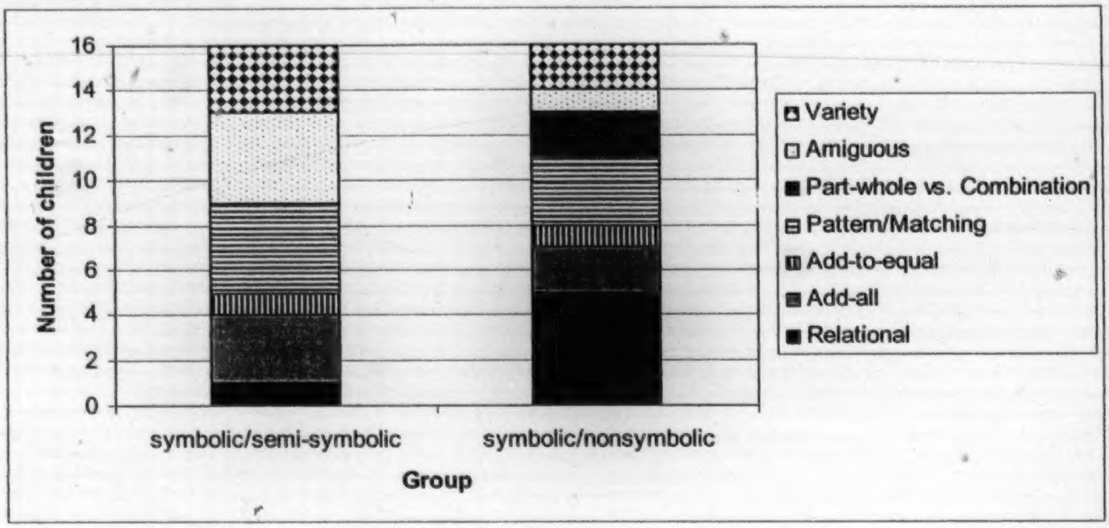


Figure 3-6. Number of children from both the symbolic/nonsymbolic (Chapter II, Study 2) and symbolic/semi-symbolic groups who used specific strategies consistently, defined as choosing the same type of solution option for at least 13 of 16 problems, on the symbolic sessions.



sign does have some impact on children's accuracy on symbolic equivalence problems (e.g., Rittle-Johnson & Alibali, 1999).

General Discussion

In this study I found that Grade 2 students can perform very differently on the same equivalence problems, dependent entirely on the condition of problem presentation and the order in which they experienced the conditions. That is, students who solved the semi-symbolic problems first had high accuracy across both the semi-symbolic and subsequent symbolic problems. Children who solved symbolic problems first, however, had accuracy no higher than would be expected with guessing on both the symbolic and subsequent semi-symbolic problems. The fact that half of the children benefited from just one session approximately a week prior to receiving symbolic problems is noteworthy, especially in light of the fact that experimenters made no direct link between the two sessions when instructing children on the tasks (just as in Chapter II, Study 2). The multiple-choice method allowed me to compare performance to chance levels, as well as to examine patterns of solution choices including whether children consistently used a particular approach for solving the equivalence problems. Examining children's strategies was useful for discovering that children who solve semi-symbolic problems first are much more likely to think about equations relationally than children who solved symbolic problems first.

Despite the similarity between Chapter II, Study 2 and the present study with respect to the nonsymbolic and semi-symbolic conditions facilitating performance on subsequent symbolic equivalence problems, an important difference between the experiments exists. That is, children in this study who solved symbolic problems first went on to have low accuracy on semi-symbolic problems. This continued difficulty on equivalence problem after experience with symbolic problems is similar to what McNeil and Alibali (2005b) refer to as change resistance. That is, poor performance on semi-symbolic problems may be due to activation of what McNeil and Alibali referred to as *knowledge of arithmetic operations*, or their *typical addition schema* (2002). Exposure to problems presented symbolically may have activated children's beliefs about the equal sign (i.e., that solutions must come after the equal sign, or that all numbers must be added and the sum put on the line), and this activation may not have been extinguished in the semi-symbolic condition in which some symbols still exist. This carry-over of poor performance was not the case in Chapter II, Study 2, in which children did well on nonsymbolic problems despite prior experience with symbolic problems or not. One explanation for children's success on nonsymbolic problems despite solving symbolic problems poorly prior is that children view nonsymbolic problems entirely separately due to the lack of symbols. This explanation does not appear likely, however, because of the evidence that children in the nonsymbolic/symbolic group appeared to learn from the nonsymbolic problems and have significantly higher accuracy on symbolic problems than children without previous nonsymbolic experience. Thus, children in this group appeared to make a connection between the two sessions and draw from their success on nonsymbolic equivalence problems and perform well on symbolic problems.

A second explanation for why children had success on nonsymbolic problems but not semi-symbolic problems in the second session is that the presence of mathematical symbols, perhaps even just the equal sign, is enough to activate children's operator beliefs. In the nonsymbolic condition, there were no mathematical symbols, and children may therefore have been able to override their inclinations to solve in an operator manner (i.e., adding up all terms without ensuring both sides of the expression are equivalent), whereas the semi-symbolic problems may have looked enough like a typical problem to evoke the operator approach to solving problems. Although the problems in the present study are not sufficient to determine exactly which symbols are (or symbol is) necessary or sufficient to hinder children's performance, it does appear that even limited symbolism can influence how children approach equivalence problems.

Clearly children's approaches to equivalence problems are influenced by prior arithmetic experiences. When children first learn to solve the complex arithmetic problems in the absence of all conventional symbols, or with just a few symbols, they perform very well. Their performance is well above chance level, and, quite surprisingly, they are able to apply their approaches to symbolic equivalence problems, which elementary students typically fail. In contrast, children without such experience often approach symbolic equivalence problems with one or more operator approaches, such as adding all terms, or adding to the equal sign. Their success on subsequent equivalence problems in other conditions depends entirely on the amount of symbolization. When subsequent problems have no conventional symbols, children's accuracy is well above chance levels. Children solve such problems relationally, understanding that the goal is to make both sides of the blue tent "the same." However, when subsequent problems contain several conventional symbols, including the equal sign, children's accuracy is lower than would be expected with random guessing.

More research is needed on the extent to which conventional symbols influence strategy use and accuracy, as well as the boundary conditions around the usefulness of using nonsymbolic or semi-symbolic conditions as a type of intervention for improving performance on symbolic equivalence problems. One important first step would be to examine children's strategies and concepts of equivalence in more detail, over a longer period of time, and with specific interventions (such as exposure to nonsymbolic, semi-symbolic, or only symbolic equivalence problems) so that performance can be tracked in relation to the interventions. In Chapter II (Study 2) and the present study, performance between just two sessions, across a one-week gap, often changed dramatically. In the next chapter I describe a microgenetic approach to studying children's view of the equal sign, adherence to addition schemas, and performance on equivalence problems, over an intensive testing period.

References

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127-145.

Baroody, A. J., & Ginsburg, H.P. (1983). The effects of instruction on children's understanding of the "equals" sign. *The Elementary School Journal, 84*, 199-212.

Behr, M., Erlwanger, S., & Nichols, E. (1980). How children view the equals sign. *Mathematics Teaching, 92*, 13-15.

Bialystok, E., & Martin, M. M. (2003). Notation to symbol: Development in children's understanding of print. *Journal of Experimental Child Psychology, 86*, 223-243.

Hughes, M. (1986). *Children and number: Difficulties in learning mathematics*. New York: Basil Blackwell.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*, 317-326.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equality and variable. *Zentralblatt für Didaktik der Mathematik – International Reviews on Mathematical Education, 37*, 1-9.

McNeil, N. M., & Alibali, M. W. (2005a). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development*, *6*, 285-306.

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466.

McNeil, N. M., & Alibali, M. W. (2002). A well-established schema can interfere with learning: The case of children's typical addition schema. In C. D. Schunn & W. Gray (Eds.), *Proceedings of the Twenty-fourth Annual Conference of the Cognitive Science Society* (pp. 661-666). Mahwah, NJ: Lawrence Erlbaum Associates.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology, 92*, 734-744.

Miller, K. (2000). Representational tools and conceptual change: The young scientist's tool kit. *Journal of Applied Developmental Psychology, 21*, 21-25.

Munn, P. (1998). Symbolic function in pre-schoolers. In C. Donlan (Ed.), The development of mathematical skills: Studies in developmental psychology (pp. 47-71). East Sussex, UK: Psychology Press.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development, 6*, 449-468.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Renwick, E. M. (1932). Children's misconceptions concerning the symbols for mathematical equality. *British Journal of Educational Psychology, 2*, 173-183.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*, 175-189.

Sáenz-Ludlow, A., & Walgamuth, C. (1998). Third graders' interpretations of equality and the equal symbol. *Educational Studies in Mathematics, 35*, 153-187.

Seo, K., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence…": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Sherman, J., & Bisanz, J. (2004, May). *Symbolic interpretation: Examining children's view of the equal sign*. Poster presented at the 16th Annual American Psychological Society, Chicago, Illinois.

CHAPTER IV

TRACKING THE PATHS OF CHANGE FROM ADHERENCE TO OPERATIONAL PATTERNS TO A RELATIONAL UNDERSTANDING OF THE EQUAL SIGN

Children's failure on symbolic equivalence problems has been well-documented in numerous studies (Alibali, 1999; Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005; McNeil & Alibali, 2005a; McNeil & Alibali, 2004; Perry, Church, & Goldin-Meadow, 1988), but in Chapters II and III, I demonstrated that performance on such problems can vary as a function of prior experience. Specifically, children who solved equivalence problems in the absence of all conventional symbols (nonsymbolic), or with just some symbols (semi-symbolic) in one session, maintained high accuracy levels when tested with symbolic problems a week later. Although the results are exciting because of the significant difference compared to children without such prior experience and the potential for interventions, I was unable to explain much about why the exposure sessions were so successful, how much exposure is necessary to inoculate children against the "confusing" symbolic context, nor any other boundary conditions surrounding change. In this chapter I explore some of the factors that make experience with equivalence problems more or less successful for inducing changes in children's performance.

Boundary Conditions of Conceptual Change

Examining boundary conditions is extremely relevant in the area of equivalence, particularly in light of the significant benefits resulting from practice with nonsymbolic problems, presented with blocks and bins. Children's hands-on experiences may have helped them map their procedures onto the symbolic problems presented a week later, yet little is known about why the blocks may have been helpful, how much exposure children needed with the blocks to achieve the benefit, and whether children could have generalized their approaches to novel equivalence problems. Therefore, exploring boundary conditions is essential for determining the effectiveness (or not) of manipulatives in mathematics instruction. Although the Grade 2 children in Chapter II maintained high accuracy on symbolic problems after experience with the blocks, some researchers caution against the assertion that manipulatives are necessarily beneficial for instruction in mathematics (Uttal, Liu, & DeLoache, 2006; Uttal, Scudder, & DeLoache, 1997). Uttal et al. (1997) noted that manipulatives are symbols unto themselves, symbolic of the quantity they are designed to represent, and are thus subject to the same misinterpretations and difficulties as other symbol systems. Before children can use the manipulatives, they need to comprehend and accept how the manipulatives represent a written symbol or concept. Although the authors do not believe that manipulatives should be banished from mathematical instruction, they caution that they should be not used as ends to themselves and viewed as the only way to teach children mathematical concepts (Uttal et al., 2006).

One method particularly useful for examining accuracy, concepts, and strategies that may change quickly is the microgenetic approach (Siegler, 2006). This approach is also useful for exploring boundary conditions that can promote or inhibit change, as researchers track children's performance closely over time. There is currently no standard for what constitutes a microgenetic design in terms of frequency of data collection or duration of examination, and methodologies differ across domains. Even within the field of developmental psychology, researchers have used very different time frames and called the time frames microgenetic. For example, Siegler and Svetina (2002) tested children's ability to complete matrices across six sessions in just three weeks, while Siegler and Jenkins (1989) administered numerous sessions over the course of 11 weeks. Although factors expected to change quickly, such as body temperature in medical research, must be measured in very short intervals (such as seconds or minutes), cognitive change is typically assessed in slightly longer intervals, such as every other day. Kwong and Varnhagen (2005), for example, tested children's spelling of new words three times a week for four to seven weeks. Because the children in Chapters II and III showed fairly convincing change during the course of one week, I wanted to monitor change more frequently, often every other day, with deliberate interventions, and for a longer duration (i.e., more sessions) that in my previous studies.

Because children's difficulty with equivalence problems is related to their adherence to operational patterns (McNeil & Alibali, 2005b), I wanted to assess children's views of the equal sign and operational patterns before and after they received the microgenetic sequence of equivalence sessions. Specifically, I wanted to examine: (a) whether children's adherence to operational patterns would change (i.e., be reduced, strengthened, or maintained) after extensive experience with equivalence problems over numerous sessions, (b) whether children's adherence to operational patterns changed more or less in relation to the condition they received equivalence problems in, such as nonsymbolic problems, and (c) whether children's adherence to operational patterns changed in relation to the number of equivalence sessions they required before solving equivalence problems correctly.

Furthermore, because children in the previous studies (Chapters II and III) demonstrated changes in performance on symbolic equivalence problems due to prior experience with equivalence problems presented in other conditions, children in this study were assigned to one of three groups: nonsymbolic, semi-symbolic, or symbolic. Children in each group received up to nine equivalence sessions, and in each equivalence session children first solved *exposure* problems presented in the condition of their assigned group (e.g., nonsymbolic), followed immediately by a worksheet with symbolic equivalence *test* problems. Finally, because the studies in Chapters II and III revealed students' propensity for relatively rapid and impressive gains in accuracy (depending on condition of the problems and the order in which they received the sessions), a two-week follow-up session was scheduled with each child to monitor whether they maintained their level of accuracy after a delay, as well as to assess children's ability to use their knowledge about equivalence to solve novel forms of equivalence problems, such as 4 + __ + 2 = 3 + 5.

The three main goals of this microgenetic study were (a) to determine the relation between children's performance on equivalence problems and their adherence to operational patterns, (b) to expand upon Study 2 (Chapter II) and the study in Chapter III by comparing children's changing performance on symbolic equivalence test problems over time in relation to exposure and practice with nonsymbolic, semi-symbolic, or symbolic only conditions, and (c) to examine the five dimensions of change (path, rate, breadth, variability, and sources of change) identified by Siegler (1996) on symbolic equivalence problems in relation to specific interventions, including children's exposure to nonsymbolic or semi-symbolic problems and feedback on exposure problems. In addition to these main goals, performance on the follow-up testing session and a clinical interview were used to evaluate children's ability to maintain accuracy, solve equivalence problems not presented during the study, and reason about conventional symbols, particularly the equal sign. Children's changing concept about equivalence, and their performance on equivalence problems, were examined in terms of accuracy on equivalence problems, the number of sessions children needed before solving symbolic equivalence problems correctly and consistently, their justifications (where useful for clarifying their strategies), confidence ratings as they solved equivalence problems, and performance on the tests of operational patterns.

Operational Patterns and Success on Equivalence Problems

The first main goal is to examine the relation between children's performance on equivalence problems and their adherence to operational patterns. Children who hold stronger addition schemas, operationalized as lower operational pattern scores, may be more resistant to change, and therefore have lower accuracy on equivalence problems that might last longer in duration (i.e., more equivalence testing sessions) than children with weaker addition schemas. That is, stronger addition schemas may result in lower accuracy overall, a longer timeframe for children to begin thinking relationally, or both. McNeil and Alibali (2005b, Studies 1 and 2) presented a good case for the change resistance theory, such that elementary students with lower operational pattern scores (i.e., stronger addition schemas), and undergraduate students whose addition schemas were activated by exposure to symbolic arithmetic problems, had worse performance on equivalence problems. Similarly, children in Chapter III who solved symbolic problems first went on to have low accuracy on semi-symbolic problems, which may have looked sufficiently like symbolic problems and therefore promoted activation of children's addition schema (i.e., solve by operating on all addends and put the answer at the end).

The microgenetic approach is useful for examining more minute details of change related to children's operational pattern scores, such as the number of strategies used and the number of sessions required to reach the stop criterion rather than just success on a single test of equivalence. Examining strategy use is important, as averaging data without considering the various strategies children display can mask the variability children demonstrate with respect to their range of approaches, the factors that influence strategy use, and the ways in which strategy use relates to change in accuracy (Siegler, 1987).

Furthermore, by administering the operational patterns test both before and after the equivalence testing sessions, I was able to examine any change in children's adherence to the addition schema as a function of their assigned condition, strategy use, as well as their performance on the testing sessions (such as the number of sessions required to meet the stop criterion). McNeil and Alibali (2002) found that children who had stronger addition schemas were less likely than children with weaker schemas to benefit from one of four one-minute interventions tailored around reconstructing equivalence problems and defining the equal sign. This microgenetic project will be the first study to examine changing schemas after repeated exposure to equivalence problems, in both the exposure conditions and symbolic test phases, for a minimum of three sessions. In contrast, other researchers, such as Rittle-Johnson (2006), have examined conceptual change after a very brief, and specifically targeted, intervention and just eight equivalence problems. The other modification that makes this study unique is the inclusion of the symbolic exposure condition. Extending the reasoning of change resistance, it can be hypothesized that children who are exposed only to symbolic problems throughout the testing sessions may be less likely to succeed on the equivalence problems and may have lower scores on the second operational patterns test than children in the other two conditions. I tested this hypothesis by comparing performance on equivalence problems and operational patterns tests by children in the three conditions.

If children in the symbolic-only group do indeed begin to reason relationally, then the change resistance theory may need to be modified. Perhaps the theory is sufficient to account for short-term or immediate deficits in performance, such as low accuracy immediately following activation of the addition schema or when children have very little experience with equivalence problems. However, the theory in its current form may not account for the potential finding that children can overcome change resistance after extended exposure to symbolic equivalence problems. Although the change resistance theory suggests that children will not achieve competence on symbolic problems after exposure to just symbolic problems (rather, that competence would be *less* likely), it might be the case that children will discover a relational strategy on their own given enough time and experience with symbolic problems. After all, most children do eventually acquire the ability to reason about the equal sign relationally in arithmetic contexts, despite many years of an operator interpretation. Perhaps it is experience with noncanonical problems, where there are terms to the right of the equal sign, that is necessary for eventually promoting relation thinking, rather than the presence or absence of symbols per se.

Effects of Prior Experience

Children's performance on equivalence problems is related to the presence or absence of conventional symbols, such that nonsymbolic problems are solved with significantly higher accuracy than the same problems presented with symbols (Chapter

II). In addition to providing direct evidence for higher success in the absence of symbols, I also discovered that children's prior experiences with equivalence problems in non- or semi-symbolic conditions had a positive effect on accuracy on subsequent symbolic problems (Chapters II and III). What is unclear from the studies in Chapters II and III, however, is whether children show immediate benefits from initial experience with problems presented in specific conditions (i.e., nonsymbolic or semi-symbolic), how long improvements in accuracy last, whether rate of improvement over time is related to views about the equal sign and operational patterns, and how experience with specific conditions influences children's ability to generalize a relational strategy to novel equivalence problems (e.g., 4 + __ + 2 = 3 + 5).

Therefore, the second main goal of this microgenetic study was to examine children's performance on symbolic equivalence problems, operational patterns, and follow-up tests in relation to their assigned condition of *exposure* problems, which were presented in one of three conditions (nonsymbolic, semi-symbolic, or symbolic) prior to all tests of symbolic equivalence performance. The main hypothesis was that children who receive experience with nonsymbolic or semi-symbolic problems prior to their symbolic equivalence tests will have higher accuracy on the symbolic problems and require fewer sessions to consistently and accurately solve symbolic problems than children who only even receive symbolic problems. Furthermore, I expected that children in the nonsymbolic and semi-symbolic groups will show improvements on measures of operational patterns tests. Assessing the effects of prior experience on adherence to operational patterns, success on symbolic equivalence problems, and ability to maintain and generalize correct strategies weeks later, have important implications for education. Should children in one or more of the groups demonstrate benefits on symbolic problems, views of the equal sign, or long-term tests of equivalence, then there would be support for allowing children to experience equivalence problems in the helpful condition(s).

Dimensions of Change

The third goal is important because children's variability and change, as found in Chapters II and III, could not be well defined due to the relatively long time frame between sessions as well as the absence of a follow-up test. Little could be said about path, such as whether children progress from failure to success directly, and how strategies might change. Similarly, in the previous studies I could not speak to the rate of change, such as whether children progress from failure to success quickly, or whether improvement takes a certain number of problems, days, or sessions. Another factor of change not addressed in Chapters II and III was breadth, namely, whether success on part-whole and combination problems is predictive of success on other types of equivalence problems. In the same sense, the previous studies did not provide much information about variability, such as whether accuracy differs between children or even within children across problems.

Finally, only minimal information could be gathered about the sources of change found in the studies in Chapters II and III. That is, improvements in performance on symbolic equivalence problems were attributed to experience with either non- or semi-symbolic problems in the previous session. What could not be addressed is whether there are other sources of change, such as feedback or repetition of instructions, that can have similar benefits to performance. Frequent observations will help me detect whether children progress through various strategies or interpretations on their way to solving equivalence problems relationally. It may be the case that for some children, or for children in particular groups, the path from operator to relational thinking is direct and quick, whereas for others there are intermediate steps (i.e., strategies and interpretations) that have not been identified before. By sampling children's strategies frequently, there was more opportunity to examine change and online reasoning than in Chapters II and III, where changes may have occurred between the two sessions and therefore were not assessed.

Each of the five dimensions were explored in relation to factors such as children's success on exposure problems, their confidence ratings, additional interventions (such as feedback), and their adherence to operational patterns. Understanding these dimensions can have practical implications in terms of designing classroom interventions or teaching techniques to prevent or correct children's misunderstanding of the equal sign. For example, learning that children typically try several strategies before solving relationally in a consistent manner and discovering that children benefit immediately from feedback are useful to educators for deciding when and how to interject children's approaches and review the relational nature of the equal sign.

Maintenance of Learning and Flexibility of Thought

Children's performance on a follow-up test was examined to determine how well children maintained performance over a two-week delay, whether maintenance related to factors such as exposure condition and performance on equivalence problems, and whether children were able to transfer their skills to novel equivalence problems. In addition to testing children's accuracy on similar and novel equivalence problems, I wanted to explore children's views about conventional symbols, including the equal sign, and the flexibility of thought they are capable of demonstrating when asked to represent specific mathematical relationships. Clinical interview sessions were used to assess the creative methods children invent for representing problems with conventional symbols or not, and when using the equal sign a money context rather than a symbolic arithmetic context. Performance on the various measures were examined with different approaches, described below, such as accuracy, number of strategies, confidence ratings, and coded verbal responses.

Method

Participants and Design

Fifty-seven Grade 2 students from two local elementary schools received an initial screening test consisting of 10 symbolic equivalence problems and five two-term (e.g., $a + b = \_\_$) and three-term (e.g., $a + \mathrm{b} + c = \_\_$) canonical arithmetic problems (Appendix E). Half of the equivalence problems were of the combination type (e.g., $a + b + c = a + \_\_$), the other half were of the part-whole type (e.g., $a + b = c + \_\_$). The screening tests were administered in March and April of the school year. Children who correctly solved at least 80% of the canonical problems but no more than 10% of the equivalence problems were considered to have met inclusion criteria. The result was a sample of 37 Grade 2 students (65% of original sample). Three Grade 2 children indicated within the first ($n = 2$) or second ($n = 1$) equivalence session that they no longer wished to participate, bringing the total sample to 34 Grade 2 children (12 males, 22 females), 11 from one school and 23 from the other. An additional two children were excluded from some analyses because they did not complete all sessions. The complete sample consisted of 32 Grade 2 students who participated in all sessions (mean age in years;months = 7;9, range: 7;2 – 9;1; one child's age was not provided by her parents), tested by one of two experimenters (17 students with one experimenter, 15 with the other).

Children received sessions in the following order: screening test, operational patterns test (first administration), equivalence testing sessions (maximum of nine sessions), clinical interview, operational patterns test (second administration), a two-week follow-up test of maintenance and transfer, and finally a classroom visit. All equivalence testing sessions were scheduled as closely together as possible considering experimenter, teacher, and student scheduling constraints. Sessions were never administered sooner than 24 hours apart (except for clinical interviews, which were often conducted immediately following the last equivalence testing session), and the equivalence testing sessions never had more than eight school days between them. The number of days between sessions was averaged across children for each session, with a range of mean intervals between equivalence testing sessions ranging from .39 to 2.32 days (Table 4-1). Children completed all sessions with one of two female experimenters, and the sessions typically were five to 20 minutes in duration. Stimuli, instructions, feedback, and problems were designed to be as similar as possible to the studies in Chapters II and III so that direct comparisons, wherever possible, could be made.

Materials and Procedures

Most sessions were videotaped in the event that experimenters needed to revisit the data for coding purposes. Incentives were not provided during the study, but all students in each of the participating classrooms received small tokens of appreciation during the classroom visits after all sessions were completed.

Screening test. Children's first session consisted of an introduction and the screening test (Appendix E) in small groups. Students were asked to solve the problems presented in the small booklet with one problem per page. They were advised that some problems might seem hard and some might seem easy, but they were asked to try their best, even if they had to guess. The booklets were handed out, and the standard instructions from Chapters II and III were used for the first two problems. No feedback or additional help was provided, although instructions were repeated once if requested. After all children had solved the first two problems, they were asked to solve the remainder of the booklet at their own pace. Once all children completed the booklet, they were told that some of them would see the experimenter again for more math activities and that others were finished with all activities.

Scores out of 10 (equivalence problems) and 5 (canonical arithmetic problems) were converted to percentages because some children missed one or two problems in their booklets. Children who performed poorly ($\leq$ 10%) on equivalence problems and well ($\geq$ 80%) on canonical problems met inclusion criteria (for participation in the remainder of the microgenetic study) because they did not begin the study with prior ability to solve equivalence problems, yet their accuracy on canonical problems indicated they were capable of the arithmetic necessary to solve equivalence problems correctly.

Table 4-1

The Number of Participating Grade 2 Students in Each Session, and the Mean Number of Intervening School Days Between Each Session and the Next

Session	Participants	Intervening days	Activities
Screening test	57	6.5	Screening test
Op. Patterns	34	8.2	Operational patterns test (1)
1	34	2.0	First equivalence testing session: Practice problems, Exposure problems: 6 equivalence problems presented in assigned condition Test problems: 9 equivalence problems presented symbolically
2	34	2.2	Second equivalence testing session: Exposure problems, Test problems
3	34	2.0	Same as previous
4	29	1.4	Same as previous
5	25	2.3	Same as previous
6	22	0.4	Same as previous
7	18	2.1	Same as previous
8	15	1.6	Same as previous
9	14	0.2	Same as previous

Interview	32	7.9	Clinical Interview
Follow-up	32	3.4	Follow-up test: Maintenance: 9 equivalence problems similar to those used in the testing sessions Follow-up: 8 novel equivalence problems
Op. Patterns	32		Operational patterns test (2)

Note: The number of children in each session decreased because some children met the *stop criterion* during the maximum number of nine sessions. The stop criterion was a minimum score of 6 (out of a maximum 9) on the symbolic testing phase for three sessions in a row. Specifically, the minimum score of 6 had to reflect at least 2 out of 3 on *each* of the three problem types used in the test phases. Two children failed to meet the stop criterion or complete all nine sessions, therefore they did not complete the clinical interview, follow-up test, or second administration of the operational patterns test.

Operational patterns. The three tasks within the operational patterns test have been used in several studies examining children's misunderstanding of the equal sign and the role of addition schemas in change resistance (McNeil & Alibali, 2002, 2005b). These tasks are designed to assess the degree to which children adhere to patterns they acquire during their experience with canonical arithmetic, particularly: (a) performing operations on all the given numbers, (b) operations occurring to the left of the equal sign and solutions going to the right of the equal sign, and (c) the equal sign signaling the total sum or the solution. The theory behind change resistance is that adherence to the above elements of the addition schema is what causes failure on equivalence problems (McNeil & Alibali, 2005b). Examining children's adherence to the three elements, and how their level of adherence relates to performance and change across sessions, is particularly important to examine in this microgenetic study because the main purpose of such an analysis is to discover the path, rate, breadth, variability, and sources of change.

Scoring the tasks in the present study differed from procedures used previously (e.g., McNeil & Alibali, 2002, 2005b). Specifically, children's responses were scored as correct or incorrect, as described below, with no further analyses of *why* children's incorrect responses were wrong. Children may have reconstructed an equivalence problem incorrectly in a manner consistent with adherence to an operational pattern, such as placing the equal sign at the end of the equation, or they may have failed to reconstruct the problem at all, perhaps signaling trouble remembering the elements of the problem. In either case, the child would receive a score of zero for the problem. Thus, scores of adherence to operational patterns in the present study may not reflect precisely the same construct as in previous studies, and should therefore be interpreted with caution.

Nonetheless, conceptual errors were carefully separated from non-conceptual errors, such as arithmetic errors. Also, individual tasks (such as reconstruction) were scored separately to identify subtests that might be predictive of change on equivalence problems.

In the first task, *equation solving*, children were asked to solve three equivalence problems (Appendix F), justify their responses, and then rate how confident they are in their solution using a seven-point scale (Appendix G). This task assesses the first element of addition schemas, performing operations on all given addends (McNeil & Alibali, 2002, 2005b). Children who adhere to this rule of operating on all addends likely responded with add-all or add-to-equal solutions, indicating that they are concerned with performing addition on all terms (or at least those to the left of the equal sign in the case of add-to-equal responses) rather than ensuring that both sides of the expression are equivalent. Solutions were scored as either correct (1 point) or incorrect (0 points), for a maximum total of 3, and a maximum average of 7 for confidence ratings.

The second set of tasks, both considered together as tests of *problem structure*, assesses the second element of addition schemas. In the first test, children were shown three equivalence problems for five seconds each and after each exposure were asked to reconstruct them on a blank piece of paper. Children's solutions were scored as incorrect (0 points) if their reconstructions contained a conceptual error such as omitting the equal sign or one of the plus signs, omitting the right addend altogether, or changing the format to "operations = answer." Children's solutions were scored as correct (1 point) if they were free of conceptual errors, even if they had reconstruction errors (such as reversing the order of the addends). The maximum score was three points. The second test is a recognition version of the first test, whereby children viewed three equivalence problems for five seconds each and after each exposure were asked to circle the matching equivalence problem from a list of seven similar-looking equations. For example, children saw "7 + 3 = 5 + __", and then saw a list with the following problems: "7 + 5 = __," "7 + 3 = 5 __," "7 + 3 + 5," "7 + 3 = __," "5 + 7 = __", "7 + 3 = 5 + __," and "7 + 3 + 5 = __", the last of which would be the likely choice for students who adhere to a typical additional schema. Children's choices were scored as correct (1 point) or incorrect (0 points) for a maximum of three points.

In the third task, *equal sign definition*, children completed two parts. In the first part they were asked to tell the experimenter what the symbol means in the context of an arithmetic equivalence problem (Appendix F). The experimenter presented the problem "5 + 3 + 3 = 2 + __", and then pointed to the equal sign. She asked the students to tell her what that symbol is and what it means, while covering both sides of the equation with her hand to leave just the equal sign exposed. Children were prompted to tell the experimenter more if they just said the symbol means "equals." Children's definitions were scored as relational (1 point), operator (0 points), or ambiguous (also 0 points). In the second part, students were told that some students from another school were asked the same question about that symbol and they had given some different answers or definitions. The experimenter asked the students to use a three-point rating scale (not so smart, kind of smart, and very smart, Appendix H) to indicate how smart the other students' definitions were. The six fictional students' "definitions" reflected operator, relational, and nonsense interpretations of the equal sign. There were three operator views, two relational views, and one nonsense view. Children's ratings were assigned points for each definition: two points for a "very smart" rating, one point for a "kind of smart" rating, and no points for a "not so smart" rating. The average rating for the four incorrect definitions was subtracted from the average rating for the two relational definitions to yield a difference score. A positive difference score is indicative of children rating the relational definitions as smarter than the less sophisticated definitions, with the maximum possible score being two points and the minimum score being negative two points.

Scores from all tasks were calculated both separately and as a composite score (ranging from negative values to a maximum score of 12 points). It is important to note that previous examinations of operational patterns scores have collapsed performance into scores of 0, 1, 2, or 3 to reflect the number of tasks in which children exhibit adherence to an addition schema (McNeil & Alibali, 2002; 2005b). A score of zero would reflect no adherence, and therefore high scores on the individual tasks. A score of 3 would reflect poor performance on all three tasks, and thus a strong adherence to the addition schema. However, due to the limited variability found in my sample, the entire score range provided more sensitivity to individual differences across and within the tasks. The tests of operational patterns were administered prior to the equivalence testing sessions and again after all sessions were completed so that children's adherence to typical addition schemas (and potential change from before to after) could be examined as a function of exposure to equivalence problems, group assignment, and the number of sessions it took children to reach the stop criterion.

Equivalence sessions. Students were pseudo-randomly assigned to one of three conditions: Nonsymbolic (n = 11), semi-symbolic (n = 12), or symbolic (n = 11). Within each participating classroom, an approximately similar number of children were assigned to each of the three conditions. Of the two children who did not complete all sessions, one was from the nonsymbolic group, the other from the semi-symbolic group. Thus, the total number of participants, available for all analyses, was 32 children, 10 in the nonsymbolic group, 11 in the semi-symbolic group, and 11 in the symbolic group. The conditions (described below) modified the presentation of exposure problems for the equivalence sessions only, and did not impact screening, operational patterns, clinical interview, or follow-up sessions. Two experimenters tested approximately half of the children within each group. The first equivalence testing session began with several practice problems, followed by an exposure phase and finally a testing phase. All other equivalence testing sessions included just the exposure and testing phases (Appendix I).

For the first equivalence testing session, children received the same practice problems as given in Chapters II (Study 2) and III: A counting string (1, 2, 3, 4, __, 6), one basic arithmetic problem (2 + 1 = __, with 2 + 2 = __ as a back-up problem), and a practice equivalence problem (1 + 1 = 1 + __). Practice problems were presented in the condition to which children were assigned, and were designed to ensure children were comfortable with the stimuli, instructions, and experimenter. Feedback was provided. Correct solutions were given 1 point, with a maximum score of three for the practice problems.

The *exposure phase* occurred at the beginning of all equivalence testing sessions and consisted of six equivalence problems presented in children's assigned condition. As in Chapters II and III, nonsymbolic problems were presented with blocks and bins, semi-symbolic problems were presented on black stands with dots representing the addends, and symbolic problems were also presented on the black stands but with all conventional mathematical symbols (Appendix A, minus the multiple-choice options). Three of the problems were part-whole, and three were combination. Children responded from an unrestricted range, rather than the multiple-choice paradigm used in Chapters II and III because I did not want to restrict the breadth of children's strategies while studying potential change over time. In the nonsymbolic condition, children solved the exposure equivalence problems by placing blocks in the empty bin. In the semi-symbolic and symbolic-only conditions, children wrote their solutions on a yellow sticky note paper that was placed above the line at the end of the problem. Between each session the note paper was removed and replaced with a fresh paper for the next child. As in the previous studies, children were asked to solve, justify their solutions, and then rate their confidence in their solutions (Appendix G). Accuracy was calculated on the six problems both as a total out of six, as well as subtotals out of three for the two problem types (Appendix I). Justifications were coded (Appendix B) and confidence scores out of 7 for each problem were recorded and averaged across the six problems. A new set of six equivalence problems was used for each exposure phase. Problems were never repeated across exposure phases, nor were they used in any of the testing phases.

Initially, no feedback was given for the exposure problems. Some children completed their sessions without ever receiving any type of feedback on the exposure problems. Some children performed poorly on the exposure problems and, subsequently, also had lower accuracy on the symbolic equivalence test problems, which were administered immediately following the exposure problems in each session. In the interest of studying change and examining the potential effects of exposure to problems presented in various conditions, a modification was implemented that involved providing feedback for children on the exposure problems. The rationale was that I could not expect to find children benefiting from exposure problems if they were not solving them correctly, or if they were not even aware they were solving them incorrectly. Therefore, I began providing feedback on the exposure problems, along with repeating the instructions once on the first test question (see below), for the remainder of the children beginning on their next session (ranging from session 4-7). Where applicable, children's accuracy, confidence, and strategy use are qualified in the Results section to reflect the onset of feedback.

Immediately following the exposure phase, children were given a worksheet containing nine symbolic equivalence problems, referred to as the *testing phase*. Children were told that they were going to solve problems similar to that which they had just done, but feedback would not be provided. Three of the problems were part-whole, three were combination, and three were a novel format designed to help me determine whether children were generalizing their understanding of the equal sign to problems they have not encountered in the exposure phases. The novel format, *three-term part-whole* (i.e., $a + b + c = d + \_\_$), were similar to both combination problems, with three terms to the left of the equal sign, and part-whole problems, with unique terms to the right of the equal sign compared to the left. Having the three types of problems allowed me to explore whether children learn to solve the problems relationally at different times, such that one type may be easier than another, as well as how well (or quickly) children transfer their experiences in the exposure phase to novel problems in the testing phase. Problems were similar in magnitude to each other and to the problems presented in the exposure phase. Children were asked to solve, justify their solutions, and then rate their confidence in their solutions for each problem. Accuracy was calculated as a total score out of nine, as well as subtotals out of three for each of the three problem types (Appendix I). Justifications were coded (Appendix B) and used to help experimenters determine children's strategies if their solutions were ambiguous. Confidence ratings were recorded and averaged across the nine problems for each session.

Children stopped the equivalence testing sessions once one of two situations occurred: they completed all nine sessions or they reached the stop criterion. The stop criterion was used because children who solved the test problems correctly and consistently did not need to complete all nine sessions. Once children consistently solved the symbolic equivalence test problems relationally in the testing phase (i.e., at least 2 out of 3 on *each* of the three problem types for three sessions in a row), they were considered to have met a convincing stop criterion. Because children responded from an unrestricted range, chance levels for reaching the stop criteria with random guessing cannot be established.

For children who received the modification of feedback during the exposure phase, the experimenter also repeated the instructions for the first question of the testing phase for each session. That is, children were given the worksheet and reminded to "put a number on the line so that when you put together these (numbers to the left of the equal sign) on this (left) side of the equal sign, you will get the same number as when you put together these (number and blank line to the right of the equal sign) on this (right) side of the equal sign." For the ten children who did not receive the feedback modification (because they had already met the stop criterion or were about to), no instructions were given in the testing phase.

Clinical interview. The interview was designed to assess children's ability to think flexibly about mathematical symbols, including the equal sign (Appendix J). The interview was typically administered immediately following children's last equivalence testing session, with a range of 0 to three school days separating the two sessions ($M =$ .22 days). Children were given a worksheet with three sections that corresponded with the three questions listed in Appendix I. The experimenter read the questions to the children and provided prompts such as "can you think of another way to show that number?" and "how could you draw this question so a friend in your class could understand it?" Children's solutions were coded to capture the amount of flexibility they demonstrated, such as fully using pictures to represent the first equation and using conventional symbols correctly in the second and third question (Appendix K).

Follow-up test. Children's ability to maintain performance on equivalence problems, as well as their ability to generalize to novel equivalence problems, was tested approximately two weeks after the clinical interview. This two-page follow-up test worksheet was administered without any instructions or feedback. Children were asked to solve the problems in the order presented as best they could, even if they had to guess. Justifications and confidence ratings were not recorded for this session. The first page contained nine symbolic equivalence problems (three each of part-whole, combination, and three-term part-whole), much like those the students had solved in their testing sessions, allowing me to test children's *maintenance*. The second page contained eight symbolic equivalence problems that were not presented in any of the other sessions, allowing me to test children's *far transfer* skills. The novel problems had blank lines to the left of the equal sign (e.g., 1 + __ + 3 = 2 + 8), blank lines immediately after the equal sign (e.g., 3 + 2 + 6 = __ + 1), and fewer or no operations (e.g., __ = 8). Tests were scored as percent correct on the first (maintenance) and second (far transfer) pages. Problems were also re-scored as correct if children responded ± 1 of the actual solution. The more lenient criterion did not affect accuracy scores, however, as they were almost identical whether scored as exactly correct or correct ± 1, and scores on the first and second page measured both ways were highly correlated (*r*(32)s > .99). Scores for exactly correct responses were used in the analyses.

Classroom visit. Once all sessions were completed, I visited each of the participating classrooms. I gave a short presentation on the equal sign and its role in mathematics. Children were allowed time to solve several equivalence problems, and the presentation ended after it appeared that all children judged noncanonical problems as being acceptable, they could solve sample equivalence problems, and they could give a relational definition of the equal sign. All children were thanked for their help (even those who did not have permission to participate in the study), and were given a pencil and sticker. Math-related books were purchased for the school libraries, and teachers were given a certificate to recognize their role and cooperation in the research project. Children were also given a worksheet to complete for fun, containing various questions about the equal sign, and a summary letter was handed out for the parents and teachers. Administrators were thanked for their participation and support, and, at one school's request, a staff presentation was given. Summary reports have been sent to the school districts and made available to interested administrators, staff, and parents.

Results and Discussion

Results are presented in an order that corresponds to the three main goals addressed in the introduction. Specifically, children's adherence to operational patterns is examined first, both in terms of original and difference scores, as well as in relation to children's experience in the equivalence sessions including group assignment and the number of sessions needed to reach the stop criterion. Effects of exposure conditions are then addressed, particularly with respect to children's performance on the symbolic equivalence test problems and their performance on the two-week follow-up test. The five dimensions of change are then described, particularly highlighting the specific paths of change children displayed across the equivalence testing sessions. Finally, results from the clinical interview are described. The number of sessions children required before reaching the stop criterion was considered the primary measure of success on symbolic equivalence test problems, although accuracy levels throughout the sessions were also used in some analyses. The number of sessions necessary to reach the stop criterion was an important measure because it is indicative of how long it took children to consistently solve the symbolic equivalence test problems correctly. Where appropriate, a distinction is made between children who participated in all nine sessions but reached the stop criterion on the last session versus those who received all nine sessions without reaching the stop criterion.

Adherence to Addition Schemas

Adherence to operational patterns (such as operating on all addends and putting the solution "at the end") has been shown to relate to failure on equivalence problems, is relatively resistant to change, may be stronger in older than younger elementary students, and has a detrimental influence on algebra performance (Knuth et al., 2005; McNeil & Alibali, 2005b). Because the operational patterns tests were administered both before and after the equivalence testing sessions, I was able to examine the change in children's adherence to operational patterns as a function of their assigned condition and their performance on the testing sessions (i.e., the number of sessions required to meet the stop criterion). McNeil and Alibali (2002) found that children who had stronger addition schemas were less likely to benefit from one of four one-minute interventions tailored around reconstructing equivalence problems and defining the equal sign than children with weaker schemas. The present study is the first to examine changing schemas after repeated exposure to equivalence problems, both exposure problems in assigned conditions and typical symbolic problems. In contrast, Rittle-Johnson (2006) tested conceptual change after a very brief, and specific, intervention and just eight equivalence problems. Children's performance on the operational patterns test was examined to address three questions: (a) how well did children do on the operational patterns test initially, (b) how well do their initial operational patterns scores predict change in equivalence performance, and (c) did children's adherence to arithmetic schemas change after the equivalence sessions?

Initial performance. In response to the first question, children's scores on the first administration of the operational patterns test were very low (*M* = 15.8%, *SD* = 13.0%), where the maximum score was 12 points. The low scores, meaning strong adherence to the addition schema, are similar to previous studies with similar age groups (i.e., McNeil & Alibali, 2002, 2005b), although a different scoring method was used. Low performance by children in the present sample is not particularly surprising considering that children were pre-screened for inclusion, with children who performed poorly on equivalence problems, and thus presumably having strong adherence to the addition schema, participating in the operational patterns test.

Children had floor levels of performance on both the equation solving ($M = 0$) and equal sign definition ($M = 0$) tasks. Surprisingly, there was no correlation between children's two problem structure tasks, with low performance on both the reconstruction ($M = 34.3\%$, $SD = 41.3\%$) and recognition ($M = 41.2\%$, $SD = 27.3\%$) components. Therefore, children's first operational patterns test demonstrate a strong adherence to the addition schema, with particularly low performance on equation solving (which they had to have in order to participate) and equal sign definitions.

Predicting change on equivalence sessions. In response to the second question, I examined children's operational patterns scores to see if adherence to an addition schema relates to their performance on equivalence problems. I hypothesized that children with stronger addition schemas would solve fewer problems correctly and may also require more trials to begin thinking relationally compared to children with weaker schemas. To test for this possibility, a multiple linear regression was used. The number of sessions children required to reach the stop criterion served as the criterion variable, and three operational patterns tasks from the first administration were used as predictors: problem structure (reconstruction), problem structure (recognition), and children's rating of other students' definitions of the equal sign. Equation solving and children's definitions of the equal sign were not included because of floor levels in performance. This set of operational patterns scores accounted for 33.4% of the variability in the number of sessions required to reach criterion ($p < .001$), with children's ability to reconstruct equivalence problems being the only significantly important predictor for the number of sessions children required to reach the stop criterion, $p = .001$. As such, the correlation between reconstruction scores and the number of sessions children completed was strong, $r(34) = -.51$, $p < .01$, whereas the relation between recognition scores and number of sessions was not, $p > .05$. As stated in the previous section, reconstruction was correlated with the total operational patterns score, but not with scores for recognition or judging other children's definitions of the equal sign.

What makes this finding particularly interesting is the fact that children's ability to correctly solve equivalence problems may be predicted by their ability to reconstruct equivalence problems after viewing each problem for just a few (five) seconds. Thus problem structure reconstruction could serve as a sort of screening test for children's readiness to learn about equivalence problems, or could indicate the amount of time and or specific help children need to become successful on symbolic equivalence problems.

Changes in adherence to the addition schema. In response to the third question, a 2 (Gender) x 3(Condition: nonsymbolic, semi-symbolic, symbolic-only) x 2(Operational patterns: first and second administration) ANOVA with repeated measures on the last variable was conducted to assess children's adherence to the addition schema before and after experiencing the equivalence testing sessions and as a function of assigned exposure condition. Children's operational pattern scores, out of a possible 12 points, did in fact improve dramatically from before ($M = 15.8\%$, $SD = 13.0\%$) to after ($M = 40.9\%$, $SD = 18.9\%$) their experience with the equivalence testing sessions, $F(1, 26) = 80.55$, $p < .001$ (Figure 4-1). Results were the same when using the 0-3 scoring range described in McNeil and Alibali (2002, 2005b), with adherence getting weaker from the first ($M =$

90.0%, *SD* = 15.2%) to second (*M* = 55.3%, *SD* = 20.0%) test, *F*(1, 26) = 80.37, $p <$ .001). The improvement from the first to the second administration of the operational patterns test did not depend on the condition children were assigned to for exposure problems or gender, however, and there were no interactions, *p*s > .05. Gender was not related to outcomes in any of the overall analyses and will not be discussed further.

Almost all children improved from the first to second administration of the operational patterns test. One child's score decreased by .25 (out of the maximum 12 points), and three other children did not change from the first to second test. All other children improved, with change scores for all 32 ranging from -.25 to 5.75 (*M* = 25.1%, *SD* = 47.9%). What makes the significant improvement on the total scores so interesting is that children's experiences between the sessions were focused on solving equivalence problems, not on changing children's conceptions about problem structure or altering their schemas for addition problems per se. The single act of trying to solve numerous equivalence problems over a fairly intensive testing period (extending for three to nine sessions with 15 problems per session) appears to have positively, albeit modestly, impacted children's overall conceptions of arithmetic structure and equivalence, as evident by changes in children's total scores. However, even at the second administration, scores were relatively low. The extremely low overall scores are suggestive of addition schemas that are somewhat resist to change, at least with minimal interventions such as experience with multiple equivalence sessions. Receiving feedback during the equivalence sessions did not impact operational patterns scores compared to children who met the stop criterion before receiving feedback, *p* >.05.

Despite overall improvements in operational patterns scores, children did not improve on all tasks equally. For instance, all children gave operator, rather than relational, definitions of the equal sign on both tests. A paired samples t-test revealed that there was no difference between children's ratings of other "students'" definitions of the equal sign either, *t*(31) < 1. Unfortunately, experience with numerous equivalence problems did not have a direct impact on children's interpretations of the equal sign in an arithmetic context. Participating in the equivalence testing sessions was enough to drastically influence children's performance on the equation solving task (with a maximum score of 3 points), however, as accuracy increased from the first (*M* = 0%, *SD* = 0%) to the second (*M* = 80.3%, *SD* = 37.7%) operational patterns task (Figure 4-2). Ninety-five percent confidence intervals for equation solving on the second test were well outside the floor-level performance on the first test. Children's view of addition problem structure was also impacted, at least on one of the two problem structure measures. Specifically, children's ability to reconstruct equivalence problems improved from before the equivalence testing sessions (*M* = 30.3%, *SD* = 39.0%) to after (*M* = 47.0%, *SD* = 44.0%), *t*(31) = -2.22, *p* < .05 (Figure 4-2). Children's ability to recognize the correct equivalence problem from a list of similar-looking addition problems did not change from before to after the equivalence testing sessions, *p* > .05.

Figure 4-1. Average total scores (in percent) on the two operational patterns tests for Grade 2 children as a function of exposure conditions (with standard error bars).

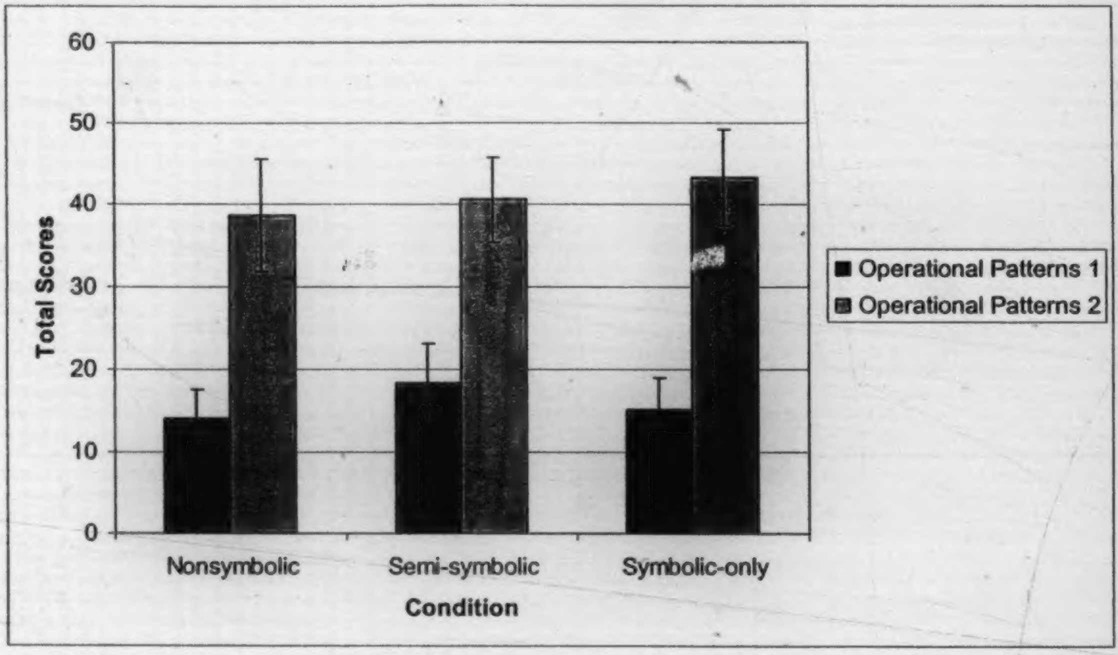
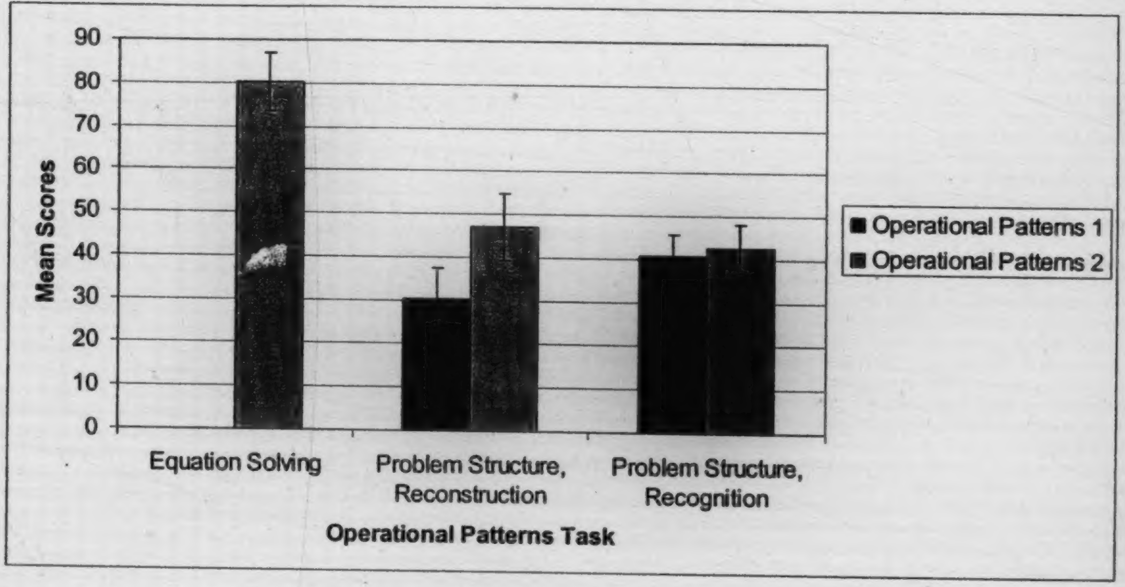


Figure 4-2. Grade 2 children's scores (in percent) on three of the operational patterns tasks from the first to second administration (with standard error bars).



Effects of Prior Experience: Relations Between Exposure Condition and Performance on Equivalence Problems

The second main goal of the study was to examine the effects of group assignment, such that children's accuracy and strategy use on exposure, equivalence, and follow-up sessions may be related to whether they experienced nonsymbolic, semi-symbolic, or symbolic problems in the exposure phases of their equivalence sessions. Based on the results described in Chapters II and III, the types of problems children encountered in the exposure phase was expected to influence performance on symbolic problems during the test phase across sessions. In particular, children who were presented with nonsymbolic or semi-symbolic problems during the exposure phase were expected to perform better on test problems than children who were exposed only to symbolic problems. Also based on the results from previous chapters, children in the nonsymbolic and semi-symbolic exposure conditions were expected to perform more accurately on exposure problems than children in the symbolic exposure conditions.

Furthermore, children's experience with equivalence problems in particular conditions might also affect performance after a relatively large time delay. For example, children who received only symbolic problems throughout the equivalence sessions may actually demonstrate higher accuracy on symbolic equivalence problems two weeks later than children in the other two conditions, as they would have had more experience with symbolic problems across the numerous sessions. Therefore, in this section, I review the influence of exposure condition (i.e., group assignment) on exposure problems, symbolic equivalence test problems (particularly the number of sessions necessary to reach the stop criterion), and the follow-up test. Effects of another experience, namely receiving feedback on exposure problems, will be examined in the *sources* of change section below.

Exposure problems. My hypothesis, based on previous studies (Chapters II and III), was that children in the nonsymbolic and semi-symbolic conditions would have higher scores on the exposure problems across the sessions, or at least on the first few sessions, than children in the symbolic condition. Because all children ($N = 34$) completed at least the first three equivalence testing sessions, a 3(Condition) x 3(Exposure Phase: Sessions 1, 2, and 3) ANOVA with repeated measures on the last variable was conducted on exposure scores for these sessions. Condition did effect children's exposure scores, $F(2, 31) = 7.48$, $p < .01$, with contrasts revealing that children in the nonsymbolic condition had higher average exposure problem scores on the first three sessions ($M = 74.8\%$) than children in the semi-symbolic ($M = 21.8\%$) or symbolic ($M = 30.3\%$) conditions. There were no other effects or interactions. Surprisingly, children who received semi-symbolic problems in the exposure phases did not have higher accuracy on exposure problems than children who received symbolic problems. Low exposure scores by children in the semi-symbolic group were unexpected based on the high accuracy by children in the previous study (Chapter III). Boundary conditions surrounding success on semi-symbolic problems are discussed below. Whether children's success on exposure problems transfers to success on symbolic test problems is examined below.

Symbolic test problems. Twenty of the Grade 2 children (58.8%) met the stop criterion at some point during the nine equivalence testing sessions; the breakdowns of students' last sessions are shown in Table 4-2. Twenty-two (68.8%) of the students began to receive feedback on the exposure problems beginning at some point beyond their third session (Table 4-3), whereas 10 children (31.3%) did not receive feedback because they had already or were about to meet the stop criterion.

Because the second goal was to assess differences in performance on symbolic equivalence problems related to the type of intervention (i.e., assigned exposure condition), I compared the last session number for children in the three conditions who met the stop criterion. Given that children who began with nonsymbolic and semi-symbolic sessions in Chapters II and III performed very well on the subsequent symbolic session, I hypothesized that children in the nonsymbolic and semi-symbolic conditions would take fewer sessions to meet the stop criterion than the children in the symbolic only condition. A one-way ANOVA on the number of sessions necessary to reach the stop criterion, however, revealed no difference among the three conditions (*M*s = 5.0 - 5.83, *SD*s = 1.62 – 2.86), $F(2, 17) < 1$. The results were similar when the last session number of all 32 children, regardless of whether they met the stop criterion, was the dependent variable (*M*s = 6.45 – 7.1, *SD*'s = 2.2 – 2.69), $F(2, 29) < 1$. There was no difference in the last session number based on experimenter (Experimenter 1: $M = 6.4$, $SD = 2.64$; Experimenter 2: $M = 7.06$, $SD = 2.14$), $F(1, 31) < 1$.

Table 4-2

Number of Grade 2 Students in Each Condition Who Met the Stop criterion, and the Last Session Completed.

		Nonsymbolic	Semi-symbolic	Symbolic
Stop Criterion?	Yes	6	7	7
	No	4	4	4
Last Session	3	2	1	2
	4	1	1	1
	5	0	1	1
	6	0	3	1
	7	1	0	2
	8	0	1	0
	9	6	4	4

Note. The children who did not complete all sessions ($n = 2$) are not represented in this table. Children who failed to meet the stop criterion completed all nine sessions.

Table 4-3

The Number of Grade 2 Children Who Received Feedback, and the Session Number on Which they Began Receiving Feedback for Each of the Three Conditions

		Nonsymbolic	Semi-symbolic	Symbolic
Feedback?	Yes	7	8	7
	No	3	3	4
Session Number	3	0	0	0
	4	2	6	3
	5	1	0	0
	6	4	2	3
	7	0	0	1

Note. The children who did not complete all sessions (n = 2) are not represented in this table. Children who failed to meet the stop criterion completed all nine sessions.

Children began to receive feedback at different points during the equivalence testing sessions, with some children receiving feedback as early as the fourth session. Because feedback was hypothesized to influence accuracy, and therefore might be confounding the condition effect analyses described above, and additional analysis was conducted to determine whether children's performance differed according to their assigned exposure conditions. A 3(Condition) x 3(Test Phases: 1, 2, 3) ANOVA with repeated-measures on the last variable revealed no difference on percent correct on symbolic test problems for the first three sessions (prior to feedback) across children in the three conditions, $F(2, 31) < 1$. Thus, children's assigned condition did not influence percent correct on initial sessions, prior to the introduction of feedback, nor to the number of sessions necessary to reach the stop criterion.

Because of the lack of a condition effect, I suggest that the benefits associated with first experiencing equivalence problems *without* conventional symbols, as found in the nonsymbolic/symbolic and semi-symbolic/symbolic groups in Chapters II and III, are subject to boundary conditions. There are several explanations for why there may not have been an effect of exposure condition. First, children may have viewed the exposure and testing phases as two separate tasks, despite the attempt made by experimenters to link the phases together by saying "now we will be solving similar problems to the ones we just did, but this time the problems will be on this worksheet." The simple act of removing the stimuli (block and bins or the black stands) and replacing it with the worksheet may have signaled a change in task that overrode the verbal attempts to link the exposure and test phases. In this study, children had already completed two very different tasks (screening test and the first operational patterns test), and interacted with one of the experimenters at least twice, before the first equivalence testing session. Thus, children may have viewed the sessions as times to do a variety of different (and unrelated) math games, limiting their ability to notice the similarities between the instructions and problem structure in the exposure and test phases. In contrast, children from the previous studies may have been attending to the similarities, such as seeing the same experimenter and going to the same testing room, rather than the differences between the tasks.

Another possibility is that the extremely short time interval between the exposure and testing phases did not allow children to consolidate what they experienced in the exposure phase and transfer their strategies or concepts to the symbolic test problems. Researchers have found that participants significantly improve their performance on certain tasks, such as visual texture discrimination, motor sequence tasks, and motor adaptation tasks, after they have slept compared to those who waited a similar amount of time without sleeping (Stickgold, 2005; Walker & Stickgold, 2006). For example, participants were trained to tap out a specific sequence with their fingers and were tested for accuracy and speed on multiple trials. When tested four to 12 hours later the same day, participants did not improve in speed, yet when tested after a night's sleep, participants could successfully tap the sequence 20% faster (Stickgold, 2005). In fact, even a 90-minute day-time nap led to significant improvements in speed compared to a similar delay when participants remained awake. Although I am not suggesting that children take a nap after being exposure to educational interventions, one might argue that by immediately testing children on the symbolic test questions, I may have missed the full potential benefit of the exposure conditions. However, children were tested over a series of days, so if sleep-dependent consolidation was an important factor, then improvements in accuracy on the symbolic test problems should have been evident beginning with the second equivalence testing session. In fact, mean performance on the symbolic test problems only increased from 21.9% (*SD* = 37.8%) to 25.6% (*SD* = 39.0%) out of nine between the first and second testing sessions.

The third possibility is that children were, in general, not experiencing sufficient levels of success on the exposure problems, thereby limiting their chance of benefiting from exposure problems. The initial procedure for the exposure phase did not require experimenters to ensure children were solving the exposure problems correctly. If children were not solving the exposure problems correctly, then there would be no reason to expect them to transfer success (since there was little success) to the symbolic test questions. Even if children were solving problems correctly, the fact that they did not know they were correct (i.e., no feedback from the experimenter) might have prevented children from trying their strategies on the symbolic test problems. The possibility of too little success on exposure problems was tested below by examining children's change in performance after the introduction of feedback part-way through the sessions.

Finally, the difference in procedure between this study and the previous studies may have limited replication of results. That is, the procedure in the microgenetic study differed from that of Chapter II (Study 2) and III in a potentially important way: No multiple-choice options were presented to restrict children's range of responses. In the previous studies, the fact that children could chose their solutions from a pre-determined (and theoretical) set of options could have restricted the range of strategies they may have otherwise used to solve the problems. The chance of children discovering a relational approach, or at least considering it as a viable option, may have been inflated in the previous studies. When asked to produce solutions on their own, children may have tried more strategies even on the exposure problems than in the previous studies. In fact, children did occasionally demonstrate a variety of strategies within a single session.

Evidence of strategy use was considered differently in this study compared to those in Chapters II and III, in which children chose responses from several clearly-defined solution options. Because solutions options were restricted in the previous studies, it was clear when there were instances of justifications matching, or not matching, the solutions. Children in this study, however, could respond from an unrestricted range, so their solutions alone were considered the primary evidence of strategy use, with justifications considered when necessary to interpret solutions that did not obviously fit a particular strategy. For example, if a child responded to 2 + 3 + 5 = 2 + __ by writing an 11 on the line, but gave an add-all justification, then the child was considered to have used an add-all strategy, with an addition error (i.e., 11 instead of 12). A response of eight, in contrast, would be considered a relational strategy. The number of different strategies children demonstrated on any given session ranged from one to five, however the overall mean of strategies on test problems per session was relatively low (*M* = 1.6. *SD* = .6).

The average number of strategies used per session by each child was calculated by first summing up the total number of different strategies each child displayed on each session (exposure and test phases separately), such as four different strategies on the first exposure phase plus three different strategies on the next exposure phase and so forth. The total number of different strategies was then divided by the number of sessions each child completed, both for exposure and test phases. A 3(Condition) x 2(Phase: Exposure, Test) ANOVA with repeated measures on the last variable revealed no difference in the mean number of strategies on exposure versus test phases, nor was there an effect of group assignment, *ps* > .05. However, an interaction between condition and phase use was present, $F(2, 31) = 8.4, p < .01$, with children in the semi-symbolic group trying more approaches on exposure problems ($M = 1.8$, $SD = .6$) than test problems ($M = 1.5$, $SD = .5$, Figure 4-3). Because children from this group had the lowest accuracy on the practice problems and were least likely to solve the practice equivalence problems correctly, it appears as though their higher number of strategies on exposure problems did not reflect an efficient, accurate, approach to solving semi-symbolic equivalence problems. An important boundary condition for the benefits of experience with nonsymbolic or semi-symbolic problems may be the presence of solution options that correspond to children's typical solutions (e.g., add-all, add-to-equal, and relational). When solving equivalence problems, children may benefit from having the number of potential strategies, and corresponding solutions, restricted to manageable and theoretical range of options.

Maintenance and transfer of equivalence skills. Very few equivalence studies have examined children's ability to maintain accuracy after a significant delay, and none have tested their success on transfer problems for the first time after a two-week interval. McNeil and Alibali (2000) tested children's ability to maintain and generalize performance on equivalence problems after a two-week follow-up test, but all problems were similar to problems previously administered in the study. The follow-up test in this study consisted of two pages, the first of which had problems of similar type to the equivalence testing sessions to test children's maintenance of performance after a two-week delay. The second page had far transfer problems, none of which resembled problems presented at any earlier point in the study.

Children who met the stop criterion would be expected to have higher performance on the follow-up test (perhaps both maintenance and far transfer problems) than children who did not meet the stop criterion, because they reached the point where they consistently and accurately solved symbolic equivalence problems during the sessions. Therefore I included the variable of reaching the stop criterion into an analysis of performance (see Table 4-2). A 3(Condition) x 2(Stop Criterion: yes, no) x 2(Follow-up Test: maintenance, far transfer) ANOVA with repeated measures on the last variable was conducted to determine whether performance on the two types of follow-up problems differed from each other and whether the difference related to children's assigned condition. Children's performance on maintenance problems ($M = 70.5\%$, $SD = 36.3$) was much higher than on far transfer problems ($M = 43.4\%$, $SD = 40.3$), $F(1, 26) = 12.6$, $p = .001$ (Figure 4-4). Children who met the stop criterion did in fact have much higher accuracy on the follow-up test ($M = 73.8\%$, $SD = 21.7\%$) than children who did not (M = 28.8%, SD = 29.8%), $F(1, 26) = 29.0$, $p < .001$.

Figure 4-3. Mean number of strategies used per session on exposure and test phases of the equivalence testing sessions for all 34 Grade 2 children (with standard error bars).

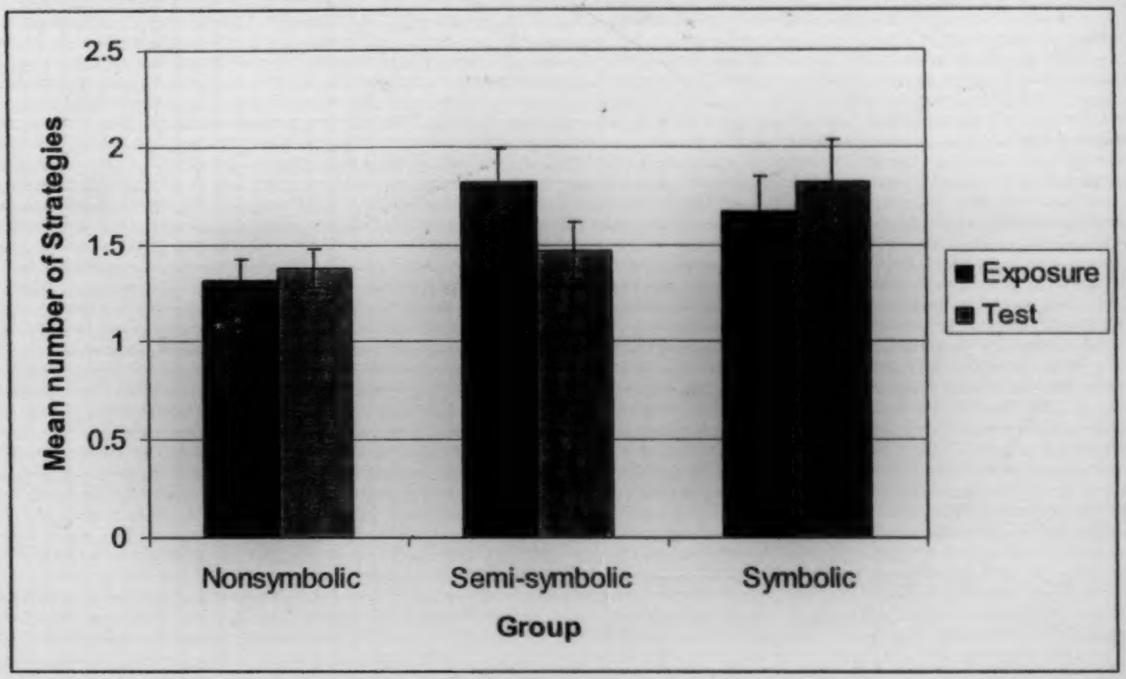
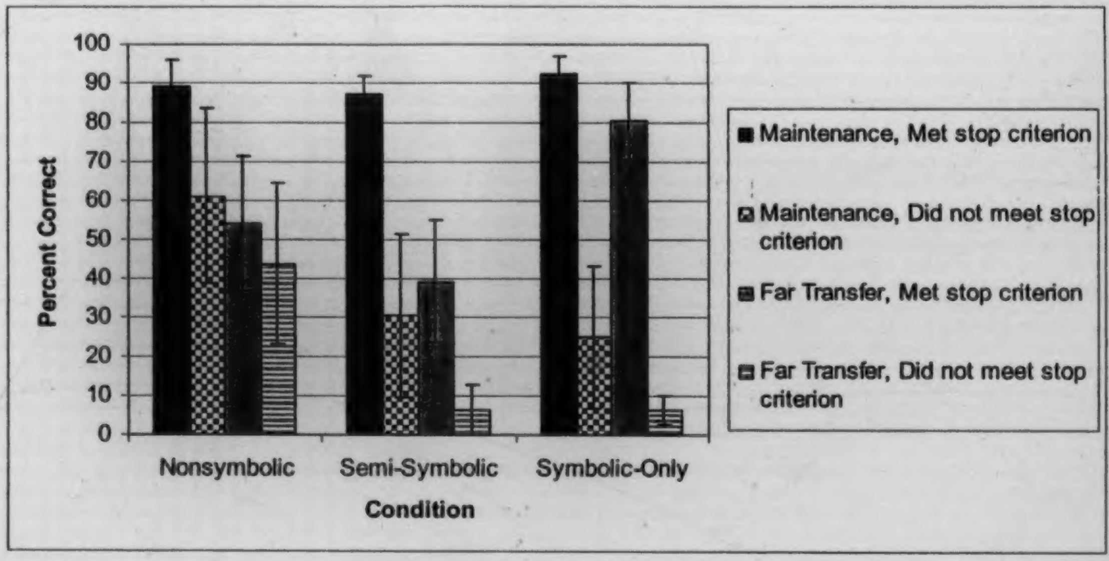


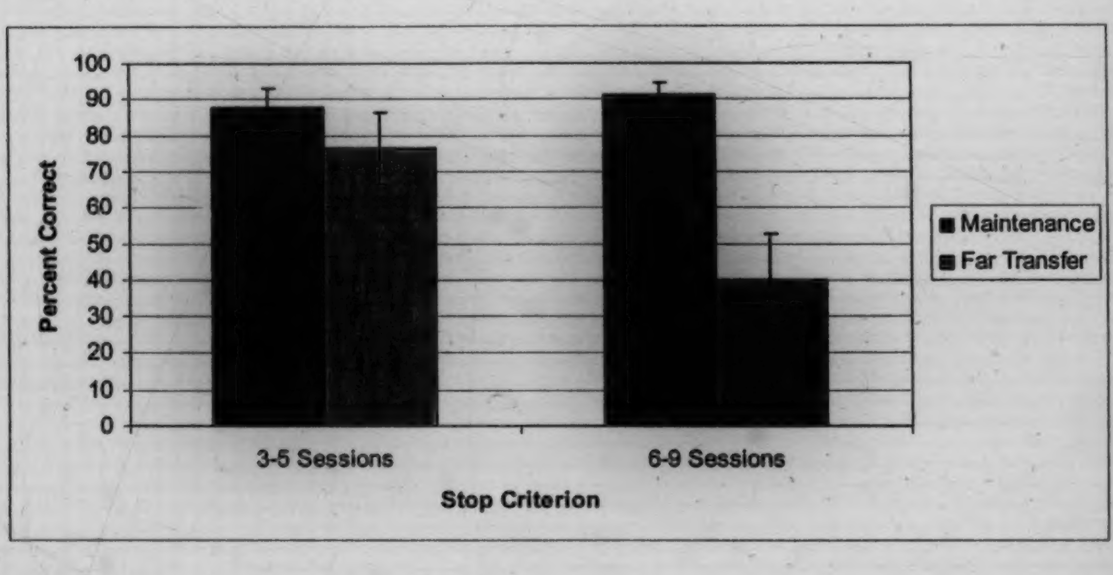
Figure 4-4. Percent correct on the maintenance and far transfer problems of the two-week follow-up test by all children ($N = 32$) who either met or did not meet the stop criterion in the three conditions (with standard error bars).



The interaction of condition and stop criterion was marginally significant, *F*(2, 26) = 3.2, *p* = .06, and deserves some attention because children in the symbolic condition appeared to perform similarly on maintenance and far transfer problems whether they met the stop criterion or not. Children who met the stop criterion from the other two conditions, however, appeared to have much higher accuracy on maintenance than far transfer problems. Perhaps experience with symbolic equivalence problems, regardless of whether experiences were successful or marked by low performance, was advantageous when having to solve new and extremely challenging far transfer problems on the follow-up test.

Even among the 20 children who met the stop criterion, there was a range with respect to the number of sessions it took for children to reach the stop criterion. Half of the students reached the stop criterion in three to five sessions, whereas the other half needed six to nine sessions to reach the stop criterion. A 2(Group by stop criterion session number: 3-5 sessions, 6-9 sessions) x 2(Follow-up Test: maintenance, far transfer) ANOVA with repeated measures on the last variable revealed that, in addition to higher accuracy on maintenance than transfer problems (as described above), there was an interaction between accuracy on the follow-up test and the grouping stop criterion variable, *F*(1, 14) = 6.7, *p* < .05 (Figure 4-5). Specifically, children who needed six to nine sessions to reach the stop criterion had lower accuracy (*M* = 40%, *SD* = 39.4%) on the far transfer problems than children who required just three to five sessions (*M* = 76.3%, *SD* = 30.9%). Thus, children who consistently and correctly solved symbolic test problems the fastest (3-5 sessions) had higher accuracy than those who required more

Figure 4-5. Performance on the two-week follow-up test by the 20 children who met the stop criterion in either three to five sessions or six to nine sessions, regardless of exposure condition (with standard error bars).



sessions when required to transfer their knowledge to transfer equivalence problems such as 3 + __ + 7 = 6 + 8.

Five Dimensions of Change

Results from the previous two sections highlight the fact that a great deal of change occurred across the various sessions. Specifically, children's adherence to operational patterns and success on symbolic equivalence problems changed, often in relation to the condition in which they received exposure problems. The third main goal of the study was to examine five dimensions of change on symbolic equivalence test problems, particularly in relation to specific interventions, including children's exposure to nonsymbolic or semi-symbolic problems and feedback on exposure problems.

Path. Examining the number of sessions needed to consistently solve symbolic equivalence problems correctly is just one way of documenting change and the potential effectiveness of an intervention. One major advantage of a microgenetic design is to track the path of children's accuracy, confidence, and strategy use across the sessions. These paths of change can be examined in relation to specific factors, such as exposure problem conditions. Furthermore, the path of children's changing accuracy on exposure problems and symbolic test problems, their confidence ratings, and the number of strategies they employ can all be tracked across the equivalence testing sessions.

Because the main goals of the present study involve children's accuracy on exposure and test problems, more so than additional measures such as confidence ratings, each child's scores on the exposure and test phases were plotted across the number of equivalence sessions they received. One rater examined the shape of the trajectories based on children's accuracy on exposure and test problems, and was able to group children into several categories. After the initial categories and criteria to describe the categories were established, a second rater, who was blind to the initial categorization, was asked to categorize children's paths of change. Inter-rater agreement was 88.2%. All differences were resolved after discussion, and can be attributed to the criteria being better defined for the second compared to the first rater. Children's paths of performance across the sessions generally fit into one of five categories (Table 4-4). These categories are described below, and in some cases, individual children's paths of change across the sessions are presented to highlight certain characteristics and variability within and across children and session. All names have been changed to guarantee anonymity.

When children improved on exposure and symbolic test problems at approximately the same time (i.e., plus or minus one session), they were considered to have *parallel improvements*. Improvements were typically impressive (e.g., 66% improvement in accuracy from one session to the next) for both exposure and test problems. Some paths of performance demonstrated *immediate success*, with high accuracy on exposure (at least four out of a maximum of six) and test (at least six out of a maximum of nine) problems on the first or second session, such that children reached the stop criterion by session three or four. Some paths of performance demonstrated *exposure before test success*, with higher scores for exposure problems than test problems (at least three or four points) at least two sessions before there were notable improvements on test problems. Some children had *low accuracy* throughout, with fairly stable but low (i.e., no more than three correct) paths of accuracy across the testing sessions, on both exposure and test problems. Finally, there were a few paths of performance that were highly *variable*, not fitting with any other category. Children with

Table 4-4

Number of children from each of the three conditions whose paths of change could be best described by one of five specific patterns.

	Nonsymbolic	Semi-symbolic	Symbolic
Parallel Improvements	2 (Joshua)	6	5
Immediate Success	3	2 (Kelly)	3
Exposure Before Test Success	4 (Tanya)	1	0
Low Accuracy	0	3 (Olivia)	1
Variable	2 (Peter)	0	2

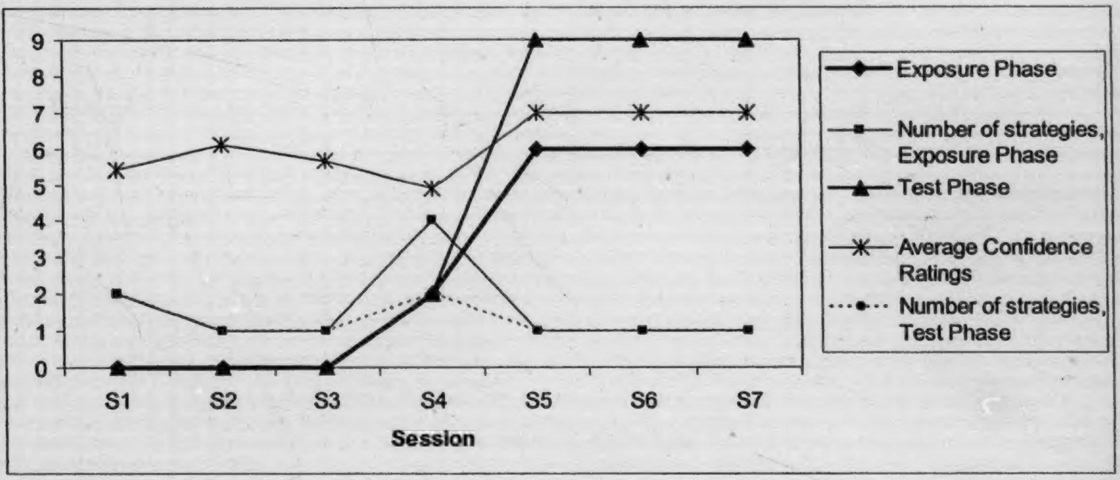
variable paths of change displayed accuracy on either exposure or test rising and falling at various, incongruous, points along the sessions, or rising dramatically once on test problems with no corresponding improvements on exposure problems.

Some characteristics about children's paths of change can be described in relation to the condition of their exposure problems (i.e., their group assignment). All children in the nonsymbolic condition solved the exposure problems perfectly by the seventh session at the latest, regardless of whether they successfully reached the stop criterion for the test problems. Three children in this group met the stop criterion for the test problems within three or four sessions, with the remaining seven children taking seven ($n = 1$) or nine ($n = 6$) sessions. Two children in the semi-symbolic group failed to have better than 50% accuracy on exposure problems on any of the sessions, and did not reach the stop criterion on the test problems. The paths to consistent success on the test problems were also more variable for children in the semi-symbolic compared to nonsymbolic condition. Children in the symbolic group also had more variable paths than those in the nonsymbolic group, with three children failing all exposure problems across all sessions, and, not surprisingly, failing to meet the stop criterion on the test problems.

Joshua is a good example of a child who had parallel improvements on exposure and test problems. He was one of the youngest children in the microgenetic study. He was eager to play the math games each day, particularly because he was in the nonsymbolic condition, which allowed him to play with blocks, but he was easily distracted by the pet turtle present in the testing room. He became so quick at solving the exposure problems that at times, after several sessions, he would just tell the experimenter the answer rather than use the blocks. He received the feedback modification on his fourth testing session, and this appears to have been pivotal for changing the way he thought about both the exposure and test problems (Figure 4-6). In fact several things happened on the fourth session as a result of being told whether his exposure problems were correct or not. Specifically, the number of strategies he employed on the exposure problems increased from one (add-all) to four (add-all, relational, add-to-equal, and ambiguous), and his accuracy on the six exposure problems increased from zero to two correct. Effects of having exposure feedback were manifested on the test problems, where Joshua's accuracy and variety of strategies increased, and his confidence ratings decreased.

Trying new approaches and being told that some of his solutions on the exposure phase were incorrect likely made Joshua feel less confident in responses on the testing phase. By the next session, he was solving all problems perfectly, his confidence was the highest possible (an average rating of seven), and he had settled on just one strategy for both exposure and test problems: using arithmetic to deliberately make both sides of the equation equal (relational strategy). When giving relational justifications for the test problems, Joshua often referred to the equal sign as the "blue tent". He exemplifies a very common pattern: Accuracy paths for exposure and test problems ran in parallel, with convincing (and immediate) benefits from receiving feedback on the exposure problems, except for one child who had already achieved high levels of accuracy on exposure and test problems before receiving feedback.

Figure 4-6. "Joshua's" paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as for his strategy use (number of different strategies) and confidence ratings (maximum = 7) across the sessions. Joshua was in the nonsymbolic condition, and received feedback beginning on the fourth session.

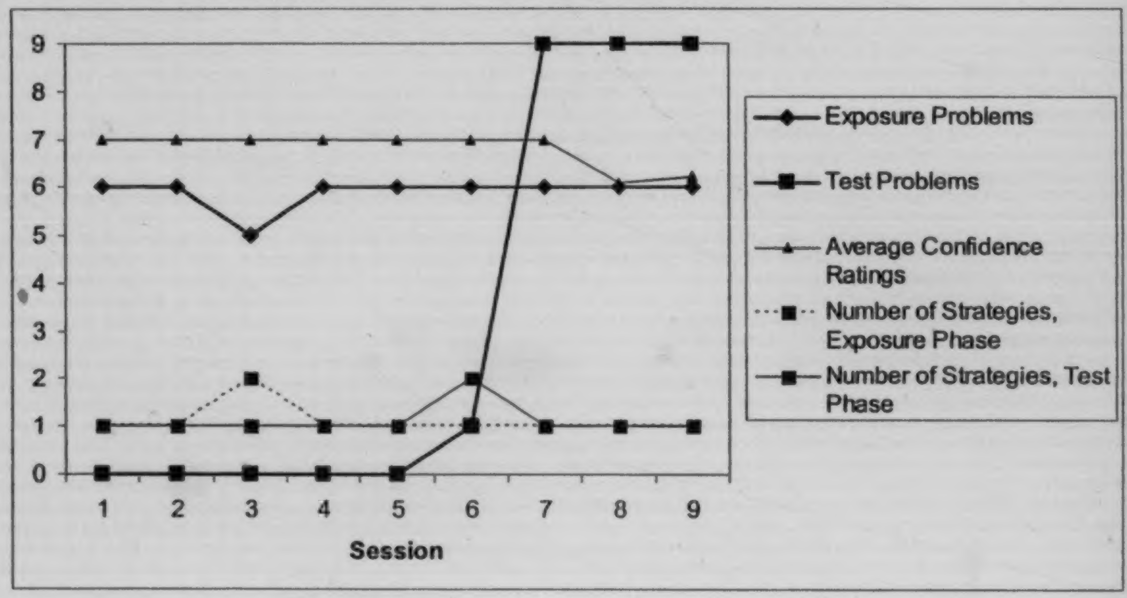


Some children appeared to improve very quickly into the sequence of equivalence testing sessions, after just one or two exposure phases. Specifically, eight children reached the stop criterion within three or four sessions, none of whom received feedback on exposure or test phases. Of these children, three were in the nonsymbolic condition, two in the semi-symbolic condition, and three in the symbolic condition. Jack was in the semi-symbolic condition, and he solved all exposure problems correctly beginning from the very first session, continuing on to solve all symbolic test problems correctly as well. His scores on the first operational pattern task were relatively high (above the 75th percentile), and his confidence on the equivalence test problems was extremely high throughout his three sessions. His accuracy and strategy use is not represented in a figure because he only had one strategy throughout all sessions (relational), and showed no variability in his accuracy. Jack serves as an example of the immediate success pattern, meeting the stop criterion within three sessions.

Tanya's total score on the first operational patterns test put her in the 29th percentile. She serves as an example of the third pattern listed above: success on exposure problems several sessions before success on symbolic test problems (Figure 4-7). Although she had perfect (or near-perfect) performance on nonsymbolic exposure problems across all sessions, the first five testing phases were plagued with a persistent, incorrect strategy, the add-all approach. Tanya's confidence ratings were very high because she assumed that adding all terms together on the worksheet was the correct approach for solving the equivalence problems. She began to receive feedback on the sixth session, but its influence on the subsequent testing phase was inconsequential because she was already solving the exposure problems correctly and confidently. The difficulty was that Tanya viewed the two phases as very separate tasks, rather than noticing the similarities between the nonsymbolic exposure problems and the symbolic test problems.

The experimenter noticed that Tanya was clearly able to solve equivalence problems correctly, as demonstrated by her impressive exposure phase scores, yet she was completely unaware that she needed to solve the symbolic problems the same way. Thus, on Tanya's last test problem on the sixth session, the experimenter repeated the standard instructions given at the beginning of the exposure phases. Tanya was immediately excited and responded: "oh, it's like the block game!" She solved the symbolic problem correctly, and proceeded to solve all remaining exposure and test problems correctly for the last three sessions. That was all it took for Tanya to benefit from her exposure to nonsymbolic problems: a tangible tie (in this case receiving the instructions again) between her experiences with the blocks and the problems presented on the worksheets. Tanya is among four other children, all from either the nonsymbolic (3) or semi-symbolic (1) condition, who solved exposure problems successfully before demonstrating high accuracy on test problems.

Figure 4-7. "Tanya's" paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as for her strategy use (number of different strategies) and confidence ratings (maximum = 7) across the sessions. Tanya was in the semi-symbolic condition, and began to receive feedback in the sixth session.



There were four children, all from the semi-symbolic (3) and symbolic (1) conditions, who never exhibited accuracy scores of three or higher on either the exposure or symbolic test problems throughout the nine sessions. Olivia began her semi-symbolic sessions successfully, solving all three practice problems correctly, including the equivalence problem. However, she solved just one exposure problem correctly, and her accuracy on both exposure and test problems was low throughout all nine sessions. She predominantly used a carry strategy, whereby she would copy a term from the left on to the blank line. She began to receive feedback on the fourth session, which appeared to lead to an increase in the variety of strategies used, but unfortunately did not improve accuracy (Figure 4-8). Olivia's paths of change highlight something common about all four children in this category, a decrease in confidence ratings over the course of the microgenetic study. Although they differed in degree and timing of confidence ratings, all four children became less sure of themselves after several sessions of low accuracy.

Not all children followed one of the four paths listed above, although most (88%) did. Peter was in the nonsymbolic condition, but he did not complete the study. His story is worth profiling, however, because of his unique path. His attention-seeking behavior during the screening test had concerned the experimenter, who worried that Peter would not be able or willing to attend throughout the numerous sessions ahead. However, once on his own (rather than in the group setting of the screening test), he appeared more calm and eager to participate. On Peter's first equivalence testing session, he had perfect accuracy on both the exposure and test phases, and his confidence was very high (Figure 4-9). He insisted on using a toy to "point" to his confidence rating on

Figure 4-8. "Olivia's" paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as her strategy confidence ratings (maximum = 7) across the sessions. Olivia was in the semi-symbolic condition, and began to receive feedback in the fourth session. Her paths of change represent low accuracy throughout the equivalence testing sessions and a decrease in confidence ratings.

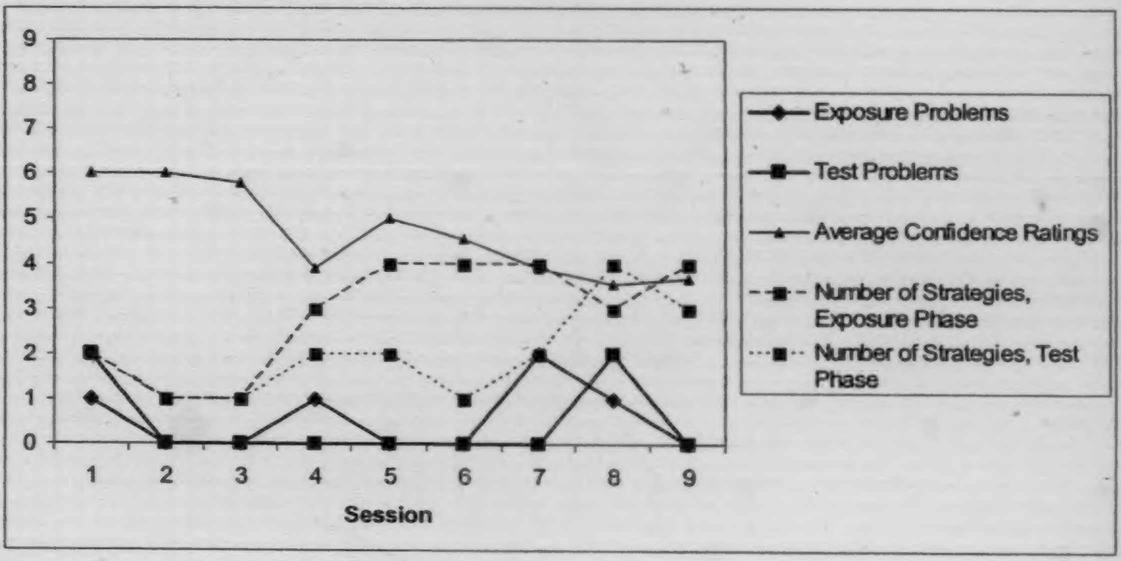
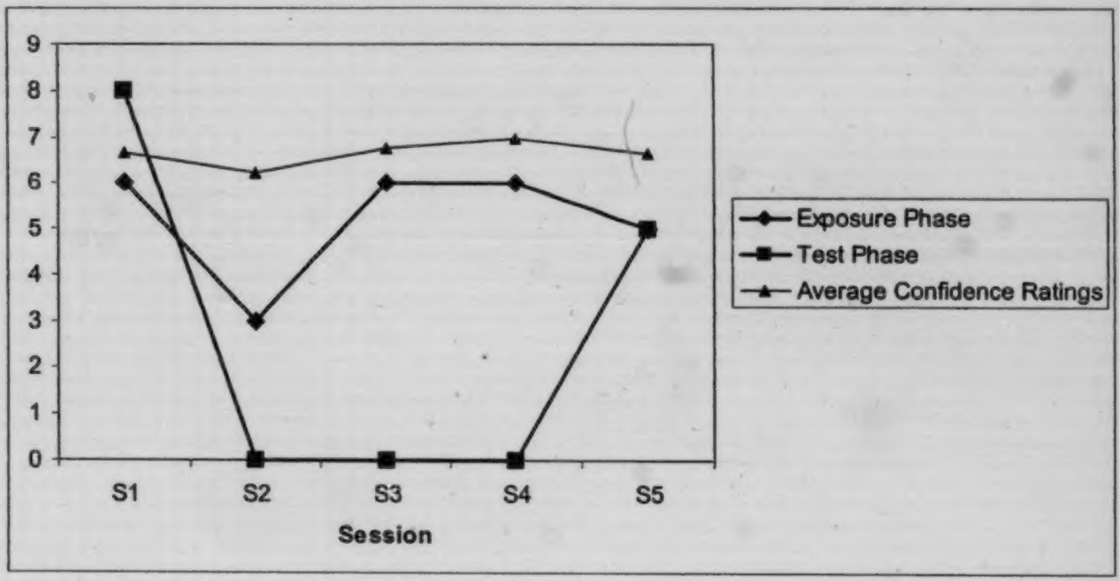


Figure 4-9. "Peter's" paths of accuracy (bold lines) on exposure (maximum = 6) and test (maximum = 9) problems, as well as his confidence ratings (maximum = 7) across the sessions. Peter was in the nonsymbolic condition, and began to receive feedback on the fifth and final session. He declined participation for any remaining sessions.



the rating scale, but did not differ from the other students in any obvious manner otherwise.

During his second testing session, a school-wide assembly was called, so Peter was dismissed to attend the pep rally in the gymnasium. The energy in the school that day was extremely high, as local media were in attendance to film and interview the students in relation to a high-profile sporting event. Peter returned to the testing room after the rally, but was noticeably fidgety and had difficulty focusing. His accuracy on the exposure problems dropped by half compared to the session before, and he failed to answer any test problems correctly. For the next two sessions, Peter solved the exposure problems correctly, but did not transfer this success to the testing phases. He was easily distracted and took a long time to respond with answers and his confidence ratings.

Peter began to receive feedback on his fifth session, with immediate and impressive effects on his test score. Although he had not solved any test problems correctly on sessions two through four, he solved five of nine correctly on the fifth session. Unfortunately Peter decided he did not wish to participate in any other sessions. The direction his accuracy and confidence paths would have taken on subsequent sessions is unknown. He is characteristic of the four children in the variable category in the sense that accuracy on either exposure or test phases, or both, appeared to fluctuate with little relation to accuracy scores on previous or later sessions, or with confidence ratings. For instance, another child with variable performance began to solve exposure problems perfectly by the fifth session, at which time she also solved six of nine test problems correctly. However, her performance on test problems had been at floor levels for the four sessions prior to the spike in accuracy, and returned to floor levels on subsequent sessions. She solved all nine symbolic test problems correctly on the ninth and final session, creating two peaks of performance along her path of accuracy on the test phases that could not be described in terms of gradual learning, prior consistency, or response to feedback.

Examining the paths of change was useful for seeing the shape of children's trajectories according to their condition, as well as for detecting patterns that emerged based on children's accuracy, confidence ratings, and strategy use. Although children in the three conditions did not differ on the number of sessions necessary to reach the stop criterion, the condition of the exposure problems did seem to influence children's paths of change somewhat (Table 4-4). Specifically, almost all children in the nonsymbolic condition fit into the first three patterns, with high accuracy on test problems occurring in tandem with improvements on exposure problems, immediately, or following several sessions with high accuracy on exposure problems only. None of the children from this group demonstrated poor accuracy throughout the duration of the testing sessions. In contrast, three of the four children who had poor performance throughout the equivalence sessions came from the semi-symbolic group. Not surprisingly, children in the symbolic group never experienced success on exposure problems before test problems, as both were presented in the same manner. The effects of condition may have acted in a subtle way, influencing children's paths of change across the sessions, and the shape of those paths, rather than directly influencing success on test problems to the extent that the stop criterion was reached at different times.

Rate. Children's paths of change (i.e., the trajectories) visually demonstrate the rate at which children changed from one session to the next. Children's biggest accuracy changes were examined to: (a) determine how quickly accuracy changed on test problems, and (b) the factors that may have affected rate of change, such as exposure condition or receiving the feedback. The transition between two sessions where the largest change in accuracy on test problems occurred (either an increase or decrease in performance) was scored as the number of problems separating the scores. For example, Joshua's test scores hovered at zero for three sessions, then he solved two test problems correctly, and the next session he solved all nine test problems correctly. The difference between his accuracy of two versus nine the next session was scored as seven. Some children were excluded from this analysis because they either did not improve or change much throughout the study (e.g., they had floor or ceiling levels of accuracy on test problems across the sessions) or their accuracy was highly variable in both directions (which was the case with four children, including Peter, who's accuracy went from extremely high to extremely low and back again).

Rate of change scores for 20 children were calculated. All 13 children whose paths of change were categorized as demonstrating parallel improvement were included, as were four of the five children from the exposure before test category. The fifth child in this category had floor levels of accuracy on the test problems across all sessions, and therefore did not demonstrate any change in accuracy. One child from the immediate success category was considered in the rate of change calculation because he had changed drastically from the first to second session, reaching the stop criterion on the fourth session. Finally, two children whose paths were categorized as variable were considered because their accuracy improved dramatically at some point during their sessions, and change was most prominent in one (i.e., improvement only) rather than two directions (increases and decreases in accuracy). The other two children from the variable category had increasing and decreasing accuracy, so rate of change scores would have been difficult to calculate, with positive and negative values that, summed together, would reflect little or no change overall.

The average score for the largest change in test scores was 6.9 points (*SD* = 1.8), or an average increase of approximately 77% from one session to the next. Therefore, for all children who showed improvement and change across the sessions (approximately 59% of all children in the study), the rate of change on test scores was surprisingly rapid, with accuracy improving drastically from one session to the next. Potential sources of change, particularly receiving feedback on exposure problems, are described in the *sources of change* section below.

Breadth. Breadth of change in this study was operationalized as children's ability to solve equivalence problems that were not presented in the exposure phase (i.e., the three-term part-whole problems), as well as their performance on the transfer problems on the follow-up test. Because performance on transfer problems was discussed above, just the results from the novel three-term problems in the testing phases are examined here. The exposure phases consisted entirely of part-whole and combination problems. The testing phases, in contrast, had both part-whole and combination problems, as well as three-term part-whole problems used to determine whether children could correctly solve equivalence problems that they had not encountered in their exposure conditions.

Accuracy on part-whole, combination, and three-term part-whole problems was examined across the equivalence testing sessions. Children's scores for each problem type were added and then divided by the total number of sessions each child completed, for a maximum average of three for each problem type. If children were successfully transferring their relational knowledge, based on practice with part-whole and combination problems, to the novel three-term part-whole problems, there should be no difference in accuracy scores on the three types of problems. A 3(Condition) x 3(Problem type: part-whole, combination, three-term part-whole) ANOVA with repeated measures on the last variable was conducted. No difference in accuracy was evident across the three problem types, $F(2, 62) = 2.38$, $p = .10$. Thus, children in all conditions were equally able to solve (or not solve) part-whole ($M = 45.3\%$, $SD = 31.3\%$), combination ($M = 46.3\%$, $SD = 30.3\%$), and three-term part-whole ($M = 43.0\%$, $SD = 32.7\%$) problems in the test phases, and scores on the three problem types were highly correlated, $r(31)$s > .94, ps < .01 (Table 4-5). The main effect of condition and the interaction were not significant.

Both combination and the novel three-term part-whole problems involved three, rather than two, addends to the left of the equal sign, making them appear slightly different (if not just longer) than typical canonical problems such as 2 + 2 = __. Because correctly solving these problems involved more calculation, and possibly required children to attend more to the structure of the problem than the part-whole problems, I wanted to see whether children's adherence to operational patterns, particularly problem structure, was related to accuracy on these two types of problems. Measures of adherence to the addition schema included the first and second administration, as well as an average measure of the two administrations, useful as a statistical index that includes both their initial adherence and that after intervention. Children's mean operational pattern score (averaged across the two test administrations) was positively correlated with accuracy on combination problems, $r(31) = .35$, $p < .05$, and on three-term part-whole problems, $r(31) = .33$, $p = .06$. Similarly, performance on the initial administration of the problem structure (reconstruction) task was correlated with accuracy on combination problems, $r(33) = .35$, $p < .05$. Considering the predictive nature of accuracy of the reconstruction task on the number of sessions necessary to reach the stop criterion as noted above, as well as the relation between scores on the first reconstruction task and success on three-term equivalence problems, this component of the operational patterns test might be particularly useful for identifying children who may be more likely to successfully solve various equivalence problems versus those who may need additional instruction and intervention.

Variability. As evident in the paths of change, children's accuracy varied considerably in at least two ways: children differed from each other in terms of accuracy, rate of change, and number of sessions necessary to reach the stop criterion (as evident by the five paths of change across the sessions); and they often varied considerably within themselves from session to session. This second type of variability was evident in many children's rapid rate of change. Peter is a good example of just how variable children's accuracy could be, progressing from high to low and back to high accuracy again across just five sessions. Variability in performance was also evident on the other measures in the microgenetic study, including the two operational patterns sessions and the two-week follow-up test, both of which were described above.

Sources. Although the primary source of change on symbolic test problems was operationalized as being the type of exposure problem condition (nonsymbolic, semi-symbolic, or symbolic only), there were other situations that may have served as a catalyst for children's changing conceptions of equivalence problems. Surprisingly, there was no overall effect of exposure condition on the number of sessions to reach the stop criterion. The introduction of feedback part way through the study, however, can be used as another opportunity to examine children's change in response to a particular event.

Change in response to feedback was examined two ways. First, children who showed evidence of change, described in the *rates* of change section, *and* received feedback (17 of the 20 children) were examined. Secondly, all children who received feedback, regardless of whether their accuracy on test problems changed, were considered.

Seventeen of the children from the *rates* of change section began to receive feedback on the exposure problems at some point during their nine equivalence testing sessions. Receiving feedback was considered to be a direct source of change when children's accuracy on the test problems improved drastically on that or the next testing session. Specifically, if feedback coincided with the two sessions used to calculate children's rate of change, and there were no other jumps in accuracy (either up or down) of a similar magnitude, then change was attributed to the feedback. Receiving the feedback modification had a profound, and immediate (i.e., within one session), effect for nine (53%) of the children, with average improvements of 7.7 points on the test problems. For seven of the other eight children, feedback began an average of 1.7 sessions before the jump in accuracy. The eighth child actually had an impressive increase in accuracy three sessions before she began to receive feedback. Children who immediately benefited from feedback could not be discriminated from those who did not benefit in terms of operational patterns scores or group assignment. One way in which these two groups of children differed was the number sessions necessary to reach the stop criterion, $F(1, 15) = 6.43$, $p < .05$, with children who immediately benefited from feedback required fewer sessions to meet the stop criterion than children who did not immediately benefit.

Twenty children received feedback, regardless of their inclusion in the *rates* of change section. To examine whether feedback might be a direct source of change, I analyzed the changes in accuracy on test problems from the session after, compared to the session before, receiving feedback. Twenty-one children completed at least one session after receiving feedback, and the average jump in accuracy was 44.4%.

There were some children who were solving exposure problems correctly before receiving feedback, and therefore feedback might not have been the factor that served as a source of change. For instance one student received feedback on the sixth session, after perfect accuracy on three exposure phases. Dianna's experimenter noticed that she was solving the nonsymbolic exposure problems correctly, but not using her skills and knowledge to solve the symbolic test problems. Her confidence ratings were extremely high, indicating that she believed she was solving the problems correctly. Much like the case with Tanya, the experimenter realized that Dianna was capable of solving equivalence problems, but was not yet noticing the similarities between the two contexts (nonsymbolic and symbolic). Deciding to see whether a small amount of prompting could change Dianna's approach, the experimenter waited while Dianna solved the first two problems of the seventh session incorrectly and then provided feedback for her incorrect solution on the second symbolic test problem. The experimenter said: "What if I told you that the answer (to 1 + 9 + 2 = 1 + __) is 11?" Dianna responded by saying: "I sometimes do not get problems right", so the experimenter repeated the standard instructions one time. Suddenly, Dianna expressed excitement as she proclaimed "now I get it!" She proceeded to solve the remaining seven problems correctly without any further help. Just being told that she had been solving the problems incorrectly, and being provided one correct solution, appeared to be influential sources of change, as Dianna solved the remaining seven of the nine test problems correctly on that session, compared to zero on all six sessions before, and had perfect accuracy on the last two sessions, meeting the stop criterion on session nine.

When the microgenetic study began, the only source of change was hypothesized to be children's exposure condition. Although the context of exposure problems may have influenced the shape of children's performance trajectories, as discussed above, there was no significant effect of condition on the number of sessions it took children to reach the stop criterion. However, results must be clarified because the presence and timing of feedback was not considered in the analysis of condition effect on number of sessions required to reach the stop criterion. Introducing the feedback modification, and carefully recording any additional interventions such as with Dianna, exposed other sources of change that, for many children, had profound, immediate effects on accuracy on the symbolic test problems.

Individual Differences

Individual differences among the 32 children who completed all the equivalence sessions were explored using correlations (Table 4-5). Performance on the screening test (equivalence and canonical problems) was not included in the analyses because of low variability (floor and ceiling levels of performance, necessary for meeting inclusion criteria). The clinical interview was not included because data were coded categorically. Children's individual differences were examined to determine whether differences in adherence to operational patterns related to performance on equivalence problems. Specifically, did overall performance on the first and second operational patterns tests relate to each other, to accuracy on the symbolic test problems, to the number of sessions children received before reaching the stop criterion, or to children's ability to maintain or transfer knowledge of equivalence?

Children's performance on the first test of adherence to typical addition patterns was positively related to the second test. Similarly, children's adherence to operational patterns, as measured by the second test, was related to children's performance on combination and the three-term part-whole symbolic test problems, as described above. Other relations existed among children's individual differences. Children's accuracy on maintenance problems related positively to their ability to solve the far transfer problems. The number of sessions children took to reach the stop criterion was negatively related to their success on the follow-up test, such that children who consistently solved the equivalence test problems correctly sooner had higher maintenance and far transfer scores than children who needed all nine sessions to reach the criterion (or who never reached the criterion), $r(31)\text{s} > .47$, $p\text{s} < .01$.

The overall pattern of children's individual differences suggest that children were likely to have similar success or failure across the three different problem types, and that success on the problem types was predictive of success on both maintenance and transfer problems presented two weeks later. Children who maintained high levels of accuracy on

Table 4-5

Correlations for the Grade 2 Children Who Completed All Equivalence Testing Sessions in the Microgenetic Study

	1	2	3	4	5	6	7	8	9	N	Mean	Std.
1. Practice Problems	1	-.20	-.15	-.14	.18	-.18	-.11	-.33	-.01	32	2.41	.56
2. Average part-whole score		1	.95**	.98**	-.93**	.16	.32	.65**	.57**	32	1.41	.94
3. Average combination score			1	.96**	-.95**	.24	.37*	.51*	.50**	32	1.45	.90
4. Average 3-term part-whole score				1	-.95**	.18	.38*	.60**	.55**	32	1.35	.98
5. Last testing session number					1	-.19	-.30	-.48**	-.53**	32	6.75	2.37
6. Operational Patterns Score (1)						1	.61**	.17	-.11	32	1.90	1.56
7. Operational Patterns Score (2)							1	.32	-.06	32	4.91	2.27
8. Follow-up test: Maintenance								1	.49**	32	70.5%	36.3%
9. Follow-up test: Transfer									1	32	43.4%	40.2%

* $p < .05$, ** $p < .01$

the three symbolic problem types were also likely to transfer their knowledge of solving various equivalence problems to novel, challenging problems.

Clinical Interview

After children completed all equivalence testing sessions, they were asked three questions to assess their ability to reason about mathematical symbols in a flexible manner. I wanted to know whether children could think about math symbols in abstract ways, such as with pictures to represent quantities, and whether they could see pictures of quantities and represent them with conventional symbols (Appendix J). Because the clinical interviews were administered after various intervening factors, including exposure to equivalence problems in particular conditions, multiple sessions with symbolic problems, and feedback, children's performance on the clinical interview are considered in relation to such factors. On the first question, the experimenter showed the children the symbolic problem 2 + 4 = 3 + __ and asked them to write or draw out the problem any way they wished, as long as they did not use any numbers or other mathematical symbols.

The first problem seemed very difficult for most of the students. Six children decided they did not even want to guess, even after several prompts, and only one of the 32 children who received the clinical interview completely represented the problem without any conventional symbols. This particular student decided to represent the equation by literally writing out the entire problem with words. She had been in the symbolic group, had reached the stop criterion in just five sessions, and did not receive feedback during the testing sessions. Her path of performance represented parallel improvements of exposure and test problems, and her scores on the second test of operational patterns was in the 75th percentile. Half of the children represented the addends with pictures (46.9%) or words (3.1%) without attempting to represent the operations or equal sign, but an additional 19% of the children attempted to represent one of the plus signs with pictures (such as circling the pictures representing the two and four to indicate putting the two groups together) but did not successfully represent the remainder of the problem. All children who attempted to represent at least one of the operator symbols were from either the nonsymbolic (3) or semi-symbolic (3) groups. Perhaps experiencing problems presented without conventional symbols in the exposure phases helped children reason about different ways to represent arithmetic expressions.

The second problem, in which children were asked to represent a word and picture problem with conventional mathematical symbols, was also difficult: Nine children (28%) did not even attempt to respond, even after numerous prompts. Thirty-four percent of the children, however, were able to represent the problem correctly, including placing the equal sign in the correct place. Six of these 11 children were from the symbolic group, two were from the nonsymbolic group, and three were from the semi-symbolic group. The high proportion (54.5%) of children from the symbolic group suggests that having experienced multiple sessions of only symbolic problems, children may be more likely to successfully represent an arithmetic expression with conventional symbols than children who experienced nonsymbolic and semi-symbolic exposure phases. All 11 children had paths of performance that were categorized as showing parallel improvements between exposure and test phases, or immediate success. Indeed, the average number of sessions before reaching the stop criterion was relatively low (M = 5.3, SD = 2.4).

Some children's responses to the second clinical interview question are telling of their persistent misunderstanding of the equal sign. Ten (31.3%) children either: (a) placed the equal sign before the last number rather than in the middle where it belonged ($n = 4$), (b) placed the equal sign at the very end of the problem ($n = 5$), or (c) used a plus sign in the place of the equal sign because there were addends to the right ($n = 1$). Unlike the 11 children who solved the problem correctly, these children required a higher number of sessions, on average, to reach the stop criterion ($M = 8.1$, $SD = 1.7$), $F(1, 19) = 9.6$, $p < .05$. The ten children were from all three groups and all five categories of the paths of change.

For the third problem, children were shown a picture with two nickels and one dime, separated by a box in which children were asked to write the appropriate symbol to show that both amounts of money were the same. Fourteen (43.8%) of children responded by writing the equal sign in the box, with another three children also responding with the equal sign after additional help and prompting, such as "can you think of a symbol from math that might go in the box to show that this side *is the same as* that side?" There were a variety of other, often creative, approaches, including: (a) putting the sum (10) in the box ($n = 3$), (b) representing the entire problem (5 + 5 = 10) ($n = 5$), and (c) drawing two arrows, pointing in opposite directions, to show the reciprocal relation between the two nickels and the dime ($n = 1$).

Children may have an easier time making the link between equations presented without conventional mathematical symbols (i.e., the second clinical interview question) and those with conventional, symbolic representations than the reverse. Eleven students solved the second problem correctly, while only 1 student correctly represented the first problem. Furthermore, the condition in which children received exposure problems may have influenced their approaches the first and second clinical interview questions. Specifically, all children who attempted to represent at least one of the operations on the first problem were from either the nonsymbolic or semi-symbolic groups. Children from these groups had experience solving arithmetic expressions that were represented either in the complete or partial absence of the conventional symbols typically used to represent the equations. Similarly, more than half of the children who correctly represented the second question with symbols were from the symbolic group.

Overall, results from the clinical interview suggest that, in general, children are not exceptionally good at thinking between two representational systems (those with and without conventional mathematical symbols) flexibly. Improving this flexibility may help children benefit from manipulatives, and other representations that do not include mathematical symbols, more quickly and efficiently. Also somewhat alarming was the fact that less than half of the Grade 2 children in this study spontaneously realized that the equal sign could be used to indicated equivalence between two sets of coins with equivalent value. Children's specific difficulty with the equal sign is also evident by their operator definitions of the symbol given in the two operational patterns tests.

General Discussion

Children's difficulty with equivalence problems has been documented in numerous studies, as has their ability to achieve some measure of success after relatively brief interventions. Less well understood was how children's adherence to operational patterns might change from one session to the next without deliberate interventions related to problem structure and the meaning of the equal sign. Furthermore, I had not explored how adherence might relate to success on equivalence problems, how children's experiences (such as nonsymbolic exposure problems or feedback) might influence progress from failure to success on equivalence problems, and what change might look like across numerous sessions, including children's paths, rates, breadth, variability, and sources of change. By examining Grade 2 children who initially failed equivalence problems over a number of sessions and weeks, I was able to address questions of changes in adherence to addition schemas, accuracy on equivalence problems, and views about conventional mathematical symbols, in relation to exposure to specific conditions, onset of feedback, and individual differences in number of sessions needed to reach the stop criterion. Analyses were conducted and presented to reflect the three main objectives of the microgenetic study: (a) to assess children's changing views about operational patterns and the equal sign, (b) to examine performance on symbolic test problems and the two-week follow-up test in relation to children's assigned exposure conditions (i.e., nonsymbolic, semi-symbolic, or symbolic), and (c) to describe children's performance, including accuracy and other indicators such as confidence ratings, across the sessions according to five dimensions of change.

Adherence to Addition Schemas

Operational patterns tests were administered both before and after the equivalence testing sessions. The pre- and post-session administrations allowed me to determine whether children's adherence to the addition schema could change after experiencing multiple sessions of equivalence problems. Furthermore, I was able to examine whether the number of sessions children needed to consistently and correctly solve equivalence problems (i.e., meeting the stop criterion) could be predicted by the strength of their adherence to operational patterns. Tasks were scored somewhat differently than described in the literature, however, with no further analyses examining whether incorrect responses were due to factors consistent with adherence to operational patterns, or to non-conceptual factors such as forgetting. Future studies will include such analyses.

In fact, children's scores on the second operational patterns test were significantly higher than on the first test, indicating that their adherence to the addition schema could be challenged and weakened indirectly (and effectively) by completing numerous sessions involving various types of equivalence problems (part-whole, combination, three-term part-whole, as well as problems presented with symbols, blocks, or semi-symbolically). This improvement was most evident on the first two tasks: equation solving and problem reconstruction. This was the first study to examine children's changing adherence without conceptual or procedural interventions designed to directly influence their operational patterns scores, and it also incorporates the longest delay between first and second administration. Solving problems with addends on both sides of the equal sign may help children realize that automatically performing operations on all addends does not guarantee success, that solutions do not always appear to the right of the equal sign, and that the answer does not always belong immediately after the equal sign.

It should be noted, however, that scores on the second administration were still relatively low (below 50%), and even though students improved remarkably in their ability to solve equivalence problems, none of them gave a relational definition of the equal sign on either operational patterns test. Although children's definitions of the equal sign and performance on equivalence problems have been shown to relate strongly (e.g., Knuth et al., 2005), the skills of solving equivalence problems versus giving verbal definitions of a symbol may develop at different rates and in response to different experiences or instruction. Finding a procedure to solve equivalence problems effectively may emerge earlier, be modified more easily, or even exist in a separate *microworld* of understanding (Lawler, 1981) than changing children's verbal definitions of the equal sign in arithmetic contexts (Rittle-Johnson & Siegler, 1998). When children have to solve novel equivalence or algebra problems, ones where their acquired procedures and addition schemas no longer work, however, a relational understanding of the equal sign is necessary to determine a viable procedure for solving. Ensuring children hold a relational view of the equal sign, then, is important for success in higher mathematics. One question that needs more research is whether solving equivalence problems correctly is a sufficient and exhaustive measure of children's relational understanding of the equal sign, or whether children's verbal definitions are a better, or just different, indicator of understanding.

Because the first administration of the operational patterns test occurred before the equivalence testing sessions, I was able to determine whether children's adherence to the addition schema could predict their success on the equivalence test problems. Again problem reconstruction emerged as an important task, a good indicator of how quickly children would reach the stop criterion during the equivalence testing sessions. This task might be a good tool for determining which children will benefit the soonest after an intervention, or, in contrast, which children need specific interventions the most. In future studies I would like to replicate the utility of this particular test for predicting children who will benefit soonest from those who will not, as well as explore why this particular task, and not the others, is especially useful.

Effects of Prior Experience

The second main goal of the microgenetic study was to examine the effect of experience in specific conditions (nonsymbolic, semi-symbolic, or symbolic) on equivalence problems. Children's accuracy on exposure phases, symbolic test problems, and the two-week follow-up test, were all examined in relation to children's assigned exposure condition. Children in the nonsymbolic group had higher accuracy on exposure problems in the first three sessions compared to children from the other two groups, suggesting that children invent a relational strategy for problems without symbols relatively quickly, compared to conditions in which some or all conventional symbols are present. High accuracy on initial nonsymbolic problems was expected, based on the previous studies (Chapters II and III), and was hypothesized to positively influence subsequent performance on symbolic equivalence problems.

Surprisingly, higher overall accuracy on exposure problems did not help children in the nonsymbolic group to reach the stop criterion on the symbolic test problems sooner than the other children. I expected that children who received nonsymbolic and semi-symbolic exposure problems during the equivalence sessions would have higher test phase accuracy, and therefore require fewer testing sessions, than children in the symbolic-only condition. Such a result would have replicated and extended results from the previous chapters. Instead of an overall effect of exposure condition, I found variability in the paths of change for children within all three conditions. Thus, there emerged perhaps the most interesting results of the microgenetic study: the existence of boundary conditions around children's ability to benefit from specific interventions (i.e., exposure conditions) and individual differences in performance across the sessions. Simply receiving equivalence problems without conventional symbols does not ensure rapid success on subsequent symbolic problems, as it appeared in Chapters II (Study 2) and III. Rather, there may exist necessary conditions, such as a specific amount of time between sessions to consolidate learning or an obvious link between equivalence problems presented in two different conditions (e.g., nonsymbolic and symbolic). Although eight children benefited from exposure problems immediately, reaching the stop criterion in three or four sessions, the vast majority took many more sessions to consistently and correctly solve the symbolic equivalence problem, with 12 children never reaching the stop criterion.

Discovering boundary conditions is extremely relevant both to cognitive developmental research and policy makers concerned with educational interventions. Uttal et al. (1997, 2006) have argued that manipulatives, such as the blocks and bins used in the nonsymbolic condition, can be a helpful way to teach children certain mathematical concepts, but there exist conditions in which using manipulatives may cause more confusion then understanding among children. They cautioned that manipulatives are symbol systems unto themselves, and as such, can add another level of representation that children must be cognizant of while solving problems with them. "Longitudinal and intensive studies of the use of manipulatives in classrooms have shown that children often fail to establish connections between manipulative sand information that the manipulatives are intended to communicate (Uttal et al., 2006, p. 176). For example, children who solved multi-digit mathematical problems with manipulatives the best actually solved much simpler problems presented with symbols with less accuracy compared to children who are not as successful with the manipulatives (Resnick & Omanson, 1987). Results from chapters II and III exemplify how manipulatives and other less symbolic systems (i.e., the semi-symbolic condition) can have impressive effects on children's success on symbolic problems presented at a later time. In this chapter, however, it is evident that there are limitations to the usefulness of these interventions that are necessary to understand before widely recommending them for classroom use.

Although additional research is necessary to further explore the boundary conditions around these interventions and others, the studies described so far shed light on some likely limitations. First of all, a definitive link between solving problems without symbols or with just some of symbols (i.e., nonsymbolic and semi-symbolic conditions), and then solving them in a symbolic context (i.e., paper-and-pencil worksheet), may be helpful to draw children's attention to the similarity between the two contexts. It may be the case that putting away the materials from the exposure problems and presenting the worksheet to the children made it appear as though the two phases involved different tasks, particularly since by the first equivalence testing session children have already participated in two sessions consisting of six different tasks (the screening test and five different operational patterns tasks). Children may have viewed this transition as the end of one "game" and the beginning of another, much like the different "games" that made up the operational patterns tests. During the operational patterns tests, children were told they would be playing five different games, and each game ended by the supplies being moved to the side and new materials appearing in front of them. This experience may have transferred to the equivalence testing session, where children again expected to play different math games with the experimenter, despite the experimenter saying that children would be solving "similar" problems.

Second, children may not have had time to consolidate what they learned, if anything, from their experience with equivalence problems in nonsymbolic and semi-symbolic conditions. In the previous studies, children were tested in two sessions presented a week apart. In the present study, children solved six exposure problems and were then immediately asked to solve nine symbolic test problems. There is evidence to suggest that some types of memory may be enhanced after a significant time interval (Stickgold, 2005), and in the present study children were tested over a series of days. So, if consolidation were an important factor, then improvements in accuracy on the symbolic test problems should have been evident beginning with the second equivalence testing session. However, few children began to show high levels of accuracy by the second session. Instead, children typically experienced a rapid rate of improvement from one session to the next, often occurring after receiving feedback.

The third potential boundary condition is success on the exposure problems. Although feedback was not provided on the equivalence problems in any of the studies, children's mean performance on the first exposure phase in the present study was 38.2%, compared to accuracy levels of approximately 70% or higher for children in the nonsymbolic/symbolic and semi-symbolic/symbolic groups on their first session in Chapters II and III. Children cannot be expected to benefit from their experiences without symbols if they were not successful in those experiences. A fourth boundary condition became evident due to a difference in procedure between the present study and the studies from Chapters II and III. That is, children who participated in the first two studies were asked to choose one of five solutions to solve each equivalence problem, one of which was always the correct (relational) option. Children in the microgenetic study, in contrast, were asked to respond from an unrestricted range, without the benefit of actually seeing the correct response in front of them. Thus, the presence of multiple choice options may be an essential boundary condition for children to learn how to solve symbolic equivalence problem from problems without symbols quickly and effectively.

Maintaining and Transferring Knowledge

Children's ability to maintain and transfer knowledge on the two-week follow-up test was also examined in relation to their assigned exposure condition group. Very few studies have examined children's accuracy on equivalence problems after a significant delay, and none have tested their ability to generalize success to novel problems for the first time on a follow-up test. In general, children had much higher accuracy on maintenance than the far transfer problems, particularly true for children in the nonsymbolic and semi-symbolic groups. Overall, children's assigned group had no influence on their success of either maintenance or transfer problems. However, children in the symbolic condition appeared to perform similarly on maintenance and far transfer problems both for children who met and those who did not meet the stop criterion, although the interaction was not significant. Perhaps mastering equivalence problems, particularly due to exposure with problems presented *only* in the symbolic condition, proved an advantage when having to solve new and extremely challenging far transfer problems on the follow-up test.

Children's success on far transfer problems was related to the number of sessions required to reach the stop criterion. Specifically, of the 20 children who reached the stop criterion, those who did so in three to five sessions had much higher accuracy on the transfer problems than those who needed six to nine sessions. What remains unclear from this study is whether it was the *duration* (i.e., few sessions versus many sessions) that was important for correctly transferring knowledge to novel problems, or some extraneous factor that influenced both the number of sessions required to reach the stop criteria and the ability to transfer understanding. Children who learned to solve equivalence problems in just three to five sessions were perhaps able to consolidate their accurate strategies more efficiently than children who required more sessions, as there were fewer occasions across which to recall strategy use.

Alternatively, children who required fewer sessions may differ on some other construct that enhances their ability to transfer knowledge. In this study, I was able to rule out constructs such as adherence to operational patterns, as no correlation existed between operational pattern scores and transfer problems, as well as simple addition skills, as all children in the study had extremely high accuracy on the canonical problems from the screening test. Several other factors have been implicated in enhancing children's ability to transfer knowledge of equivalence problems, including specific types of instructions (e.g., procedural versus conceptual, Rittle-Johnson & Alibali, 1999), self-explanation (Rittle-Johnson, 2006), but these factors were not manipulated in the present study.

Five Dimensions of Change

Examining children over numerous sessions allowed me to examine their paths, rate, breadth, variability, and sources of change, which was the third main goal of the study. I was able to identify five basic paths of change for children from across the three conditions, demonstrating that children did not merely change from failure to success on symbolic equivalence problems in a linear fashion. Most children did progress from low to high accuracy over a number of sessions, but there was wide variability about the number of sessions necessary to demonstrate improvement on both exposure and test problems.

This is the first study to document children's paths of change from failure to success on equivalence problems over numerous sessions, and the first to characterize specific patterns. Children experience with exposure problems in particular conditions may have influenced children's paths of change, at least to some degree. For example, none of the children from the nonsymbolic group could be categorized as having low accuracy throughout all sessions. When children did improve on symbolic test problems, which was the case for many of the students, their rate of change was often extremely rapid, with accuracy improving 77% from one session to the next. This rapid improvement often occurred immediately or very soon (less than two sessions) after beginning to receive feedback part-way through the study. Discovering that they were or were not solving the exposure problems correctly had a profound impact on some children's accuracy on both exposure and test problems.

Children demonstrated breadth of understanding equivalence problems in two ways. First, there was no overall difference in accuracy for the three problem types (part-whole, combination, and three-term part-whole) in the symbolic phases. Despite children in the present study having similar accuracy on the three problem types, both combination and three-term part-whole problems had three addends to the left of the equal sign, and these two problem types both correlated with two important factors: Adherence to operational patterns, as measured by children's average operational pattern scores across the two administrations; and children's ability to maintain and transfer knowledge of symbolic equivalence problems, as measured by the follow-up test. Therefore, these two problem types may be more indicative of children's views about the equal sign and their ability to maintain and transfer a relational strategy to novel problems compared to part-whole problems.

The second way in which children demonstrated impressive breadth of understanding, not unrelated to the first way, was their ability to transfer their relational strategy for solving problems to the transfer problems on the two-week follow-up test. Transfer problems looked very different from the three problem types used throughout the equivalence testing sessions, yet some children demonstrated an impressive ability to transfer their relational strategy to the challenging problems, despite the two-week interval. Thus, there was a great deal of variability between and within children in terms of accuracy, number of sessions required to meet the stop criterion, paths of change, adherence to operational patterns, and ability to maintain and transfer knowledge.

Representing Mathematical Relations: Thinking Flexibly About Symbols

In addition to tracking children's changing views of operational patterns and the equal sign, I was interested in assessing children's ability to think about arithmetic equations and relations with conventional and non-conventional mathematical symbol systems. During the clinical interview, children were asked to solve three problems. On the first question, a symbolic equation was given and children were asked to represent the problem any way they wanted as long as they did not use conventional symbols. Only one of 32 children was able to represent the entire problem, using written notation rather than mathematical symbols. Many children attempted to represent at least part of the problem, including at least one of the addition signs, and all who were successful to that end had been in the nonsymbolic group. Perhaps experience with arithmetic equations presented nonsymbolically helped children reason about alternative representational formats, such as drawing icons to represent the addends, and circling groups of icons to illustrate addition.

On the second question, children were shown an equation represented without any conventional symbols and invited to represent the problem with typical mathematical symbols. Although nine children did not even want to attempt the problem, eleven of the children represented the problem correctly using symbols. Ten children demonstrated a misunderstanding of the equal sign when attempting to represent the problem with symbols, placing the equal sign at the very end of the expression, or directly before the last addend, rather than in the middle to represent "is the same as." These ten children needed significantly more sessions to reach the stop criterion than children who had correctly represented the problems, suggesting that being able to flexibly reason about different representational systems might be related to children's ability to learn a relational strategy for equivalence problems more quickly.

On the third question, children were asked to think of a symbol that would show that the amount represented by two nickels was the same as one dime. Fewer than half of the students spontaneously wrote the equal sign between the two sides of the money comparison. Greeno and Hall (1997) described how representations should be taught as tools, not as ends in themselves. These "tools" are useful for communicating ideas and concepts, reasoning about particular problems, and constructing hypothetical models. Children would be greatly served if they understood early on that the equal sign represents equality in all contexts, and can be used to express equal relations in a variety of contexts, including math, money, measurement, and social or cultural judgments.

Summary

In summary, the Grade 2 children in the microgenetic study showed great diversity from each other as well as great diversity within themselves from one session to the next. Important boundary conditions appear to exist that heavily influence the potential success of specific interventions. Determining these conditions must not be ignored in either cognitive developmental or educational research, as children's ability to draw success and conceptual change is clearly related to minute, but important, details within an intervention. Studying the limitations and exploring the classroom conditions becomes crucial for understanding the mechanisms of conceptual change and proposing instructional reform.

A potentially vital condition within the classroom is the extent to which teachers are focusing instruction on a relational understanding of the equal sign in mathematical contexts. Examining teachers' perceptions of children's understanding of the equal sign, the instructional techniques they employ to encourage a relational view, and the strategies they believe students use, is essential for understanding how some students show convincing capabilities on equivalence problems while others overwhelmingly fail, and how students learn to map their experiences of equivalence in general, such as with a teeter-totter, onto arithmetic expressions. In the next chapter I explore teachers' perceptions in order to begin addressing these important questions.

References

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127-145.

Greeno, J. G., & Hall, R. P. (1997). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan, 78*, 361-367.

Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: Equality and variable. *Zentralblatt für Didaktik der Mathematik – International Reviews on Mathematical Education, 37*, 1-9.

Kwong, T. E., & Varnhagen, C. K. (2005). Strategy development and learning to spell new words: Generalization of a process. *Developmental Psychology, 41*, 148-159.

McNeil, N. M., & Alibali, M. W. (2005a). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development, 6*, 285-306.

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466.

McNeil, N. M., & Alibali, M. W. (2002). A well-established schema can interfere with learning: The case of children's typical addition schema. In C. D. Schunn & W. Gray (Eds.), *Proceedings of the Twenty-fourth Annual Conference of the*

Cognitive Science Society (pp. 661-666). Mahwah, NJ: Lawrence Erlbaum Associates.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology, 92*, 734-744.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Resnick, L. B., & Omanson, S. F. (1987). Learning to understand arithmetic. *Advances in Instructional Psychology, 3*, 41-96.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77*, 1-15.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*, 175-189.

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills: Studies in developmental psychology* (pp. 75-110). Hove, UK: Psychology Press.

Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon and R. M. Lerner (Series Eds.), and D. Kuhn and R. S. Siegler (Vol. Eds.), *Handbook of Child Psychology: Volume 2: Cognition, Perception, and Language* (6th ed., pp. 464-510). Hoboken, NJ: Wiley.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*, 250-264.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Siegler, R. S., & Svetina, M. (2002). A microgenetic/cross-sectional study of matrix completion: Comparing short-term and long-term change. *Child Development, 73*, 793-809.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature, 437*, 1272-1278.

Uttal, D. H., Liu, L. L., & DeLoache, J. S. (2006). Concreteness and symbolic development. In L. Balter and C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues, 2nd Ed.* (pp. 167-184). New York: Psychology Press.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology, 18*, 37-54.

Walker, M. P., & Stickgold, R. (2006). Sleep, memory, and plasticity. *Annual Reviews in Psychology, 57*, 139-166.

CHAPTER V

TEACHERS' PERCEPTIONS ABOUT CHILDRENS' UNDERSTANDING OF THE EQUAL SIGN

Despite the overwhelming evidence to suggest that elementary students do not have a relational understanding of the equal sign in symbolic arithmetic contexts, including evidence from the two operational patterns tasks described above, an assumption exists that teachers are typically not aware of this difficulty (Seo & Ginsburg, 2003). One reason for teachers' possible ignorance about this phenomenon is that students generally receive canonical arithmetic problems in school, problems with which an operator interpretation of the equal sign is sufficient for solving correctly. That is, children who hold the addition schema, including the operator interpretation of the equal sign, believe that the answer always follows the equal sign, to the right of the problem. With canonical problems such as 3 + 4 = __, this belief is sufficient for rapid and consistent success. Speed drills typically consist of canonical problems, and mathematical textbooks tend to present a great majority of problems in this format, further encouraging an operator view of the equal sign or, at the very least, failing to challenge children's operator interpretations sufficiently (Seo & Ginsburg, 2003).

Even teachers who intentionally instruct children about the equal sign in various contexts are often satisfied that children have a fully relational, generalized, understanding of the equal sign because of relational responses in contexts such as the comparison of coins (e.g., 5 pennies = 1 nickel) or language arts (e.g., "little = small"). Unfortunately, relational responses in contexts other than arithmetic cannot be used as evidence of a relational understanding of the equal sign in symbolic arithmetic contexts.

Instead, children appear to have two interpretations that are not integrated and are strongly influenced by context (Baroody & Ginsburg, 1983; Seo & Ginsburg, 2003). If teachers are not aware of this misunderstanding, then they likely will not be equipped to help expand and integrate children's interpretations of the equal sign.

Link Between Teachers' Perceptions and Students' Performance

Little is known about teachers' views of student performance on equivalence problems, however, or how teachers actually instruct students about the equal sign in the classroom. There is a relation between teachers' perceptions and children's performance, making an examination of what teachers' perceive about students' equivalence understanding very important for determining how to correct and prevent problems (Asquith, Stephens, Knuth, & Alibali, 2006; Nathan & Koedinger, 2000a; Stephens, 2006). "The beliefs that teachers hold are so instrumental in shaping mathematics teachers' decisions and actions, [that] it is important that these beliefs be a focus of educational research; in addition, because these decisions affect students' learning experiences so directly, it is equally important to understand the accuracy of teachers' beliefs" (Nathan & Koedinger, 2000a, p. 210). Misperception of children's concepts of equivalence, for example, may result in teachers failing to focus on experiences or principles that could help students view the equal sign relationally in symbolic contexts.

Teachers may develop an "expert blindspot" when it comes to teaching students about the equal sign, an idea developed by Nathan and Petrosino (2003) to explain students' well-documented difficulty with arithmetic and algebra. Teachers with advanced knowledge of arithmetic, for example, may misattribute children's difficulty with equivalence problems to a deficit in equation mastery, rather than a misunderstanding of the equal sign. Excessive practice on canonical arithmetic, such as "drill and kill" or speeded exercises administered with the hope of improving equation mastery, may indeed have the reverse effect (English & Halford, 1995), as predicted by the change resistance account (McNeil & Alibali, 2005b). Furthermore, teachers may overestimate children's mastery of mathematical symbols because of the focus that textbooks and curriculum have in the classroom. Because children are typically exposed first to symbols and then to applications of the concepts behind those symbols, teachers may hold a symbol-precedence view (Nathan & Koedinger, 2000a, 2000b). Thus, teachers may spend less time than necessary teaching and reminding students about the meaning of mathematical symbols, assuming instead that children already understand and can use them appropriately.

Asquith et al. (2006) interviewed 20 middle-school (Grades 6-8) teachers to explore their perceptions of students' concepts of both the equal sign and variable, two core aspects of algebra. Teachers were shown copies of a worksheet that Grade 6 – 8 students completed (in which students were asked to define the equal sign, determine whether two equations were equivalent, define variable, and decide which algebraic expression was larger), and asked three questions: (a) what answers (correct or not) would you expect your students to give and what strategies might they have used, (b) how many students from across the district would get this problem correct and how many would use each strategy, and (c) explain your reasoning for your answers. In response to the second question, teachers tended to overestimate students' understanding of the equal sign. Although students typically progressed from an operator to a relational view of the equal sign, they did not do so as quickly as their teachers predicted. For example, teachers predicted that 73% of the Grade 7 students would give a relational interpretation of the equal sign, yet only 37% actually did so.

Asquith et al. (2006) also noted that teachers may not be aware of the relation between students' understanding of the equal sign and their performance on algebraic problems such as judging the equivalence of different equations. In response to the third question, in which teachers were asked to explain why they thought students would respond in a particular way (correctly or incorrectly), teachers' responses tended to reflect one of three reasons: student characteristics (such as motivational factors, cognitive ability, and experiences with notation), curriculum (i.e., not enough focus on a particular concept or symbol), or teachers' own classroom observations and instructional practices (i.e., "they learned about this in the first grade"). I will examine the responses of teachers in the present study to determine whether teachers' perceptions fall into similar categories.

In this study I explored teachers' beliefs about children's view of the equal sign, the ways in which teachers instruct students about equivalence, and how well they think children do on tests of equivalence. The goal was to describe the instructional approaches currently used in several elementary classrooms, and the beliefs that guide instruction, so that children's accuracy and strategies can be put into a larger context. For example, discovering that teachers are unaware of children's misunderstanding and that children are often presented with canonical problems would help strengthen the view that children fail to map their understanding of equivalence to symbolic contexts because of their experiences in the classroom. Finding that teachers are often unaware and often use canonical examples would in turn provide support for the argument that teachers need to be made aware of this common difficulty. In contrast, if teachers appear aware of children's misunderstanding and are employing techniques designed to encourage a relational view of the equal sign, this would provide evidence for the argument that students require additional supports to help them adapt a fully relational view of the equal sign in mathematical contexts.

Method

Elementary school teachers (Grades 1-6) from five local schools were invited to participate in completing a short questionnaire (Appendix L). Two of the five schools had participated in the microgenetic study (Chapter IV). At one of those two schools, I visited each elementary classroom to describe the study and solicit participation, with five teachers responding positively. At the other school I gave a short presentation at a staff meeting, with four teachers agreeing to participate. Principals at the three schools which were not represented in the microgenetic study asked if they could approach the teachers themselves, resulting in a total of 8 additional teachers.

Of the 17 participating teachers, five taught Grade 1 (or Grade 1-2 split), two taught Grade 2, three taught Grade 3, three taught Grade 4 (or Grades 4-5 split), and four taught Grades 5 and/or 6. For some analyses, grades were combined in the following ways: Group 1 (Grades 1 or 1-2 split, $n = 5$), Group 2 (Grades 2 or 3, $n = 5$) and Group 3 (Grades 4, 5, and 6, or split classes in this range, $n = 7$). After giving consent to participate, teachers completed the survey and sealed them in an envelope. One teacher (Grade 1) completed the survey with me in a one-on-one interview.

Results and Discussion

Responses to the individual questions on the surveys were transcribed verbatim, and then coded to reflect themes that addressed questions such as how teachers typically instruct students about equivalence, and how they believe students solve equivalence problems. A second rater also coded teacher responses, with inter-rater reliability across the survey questions being 85.9%. All differences in coding were resolved through discussion. Explanations and rationales for how each of the ten teacher questionnaire items was coded are discussed individually below and organized to address three specific questions: How do teachers perceive students' understanding of the equal sign? How do teachers instruct students about the equal sign? How well do teachers think students will perform on tests of equivalence?

Teachers' Perceptions of Children's View of the Equal Sign

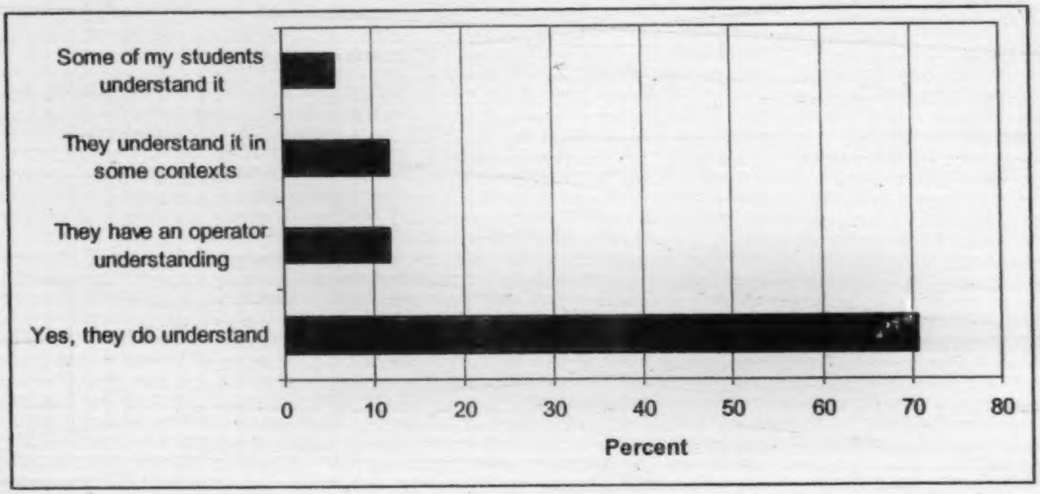
Discovering teachers' perceptions about children's understanding of the equal sign is imperative for putting children's (mis)conceptions into context. If teachers are unaware of students misunderstanding, then students are unlikely to benefit from deliberate instruction on the relational meaning of the symbol in the classroom. In the first question, teachers were asked to comment on whether they believed students in their class had a good understanding of the equal sign and how they came to that judgment about students' beliefs. As shown in Figure 5-1, most teachers (70.6%) commented that their students did indeed understand the equal sign, typically without expanding further on why they believed their students knew about the symbol. Two teachers each responded that their students had an operator understanding only or that their understanding depended on the context of the equal sign. One, for example, said:

> I think many of the Grade 3s believe the equal sign "points to" the answer. Having approached a variety of problem solving practice questions, many of the students' responses do not indicate an awareness of the equality of the 2 sides or a balance. Some students, however, seem to understand this idea quickly, once taught.

Another teacher stated that some of her students understood the symbol while others in her class did not. Even among the teachers with the youngest students (Group 1), 60% believed that their students have a fairly good understanding of what the equal sign means. For teachers with the oldest students (Group 3), 85.7% believed that their students have a good understanding of the symbol. These high estimations run in sharp contrast to data found in Chapters II, III and IV, as well as in previous research, that suggests very few elementary school children have a relational understanding of the equal sign, particularly in arithmetic contexts (e.g., Baroody & Ginsburg, 1983; Sáenz-Ludlow & Walgamuth, 1998; Seo & Ginsburg, 2003).

All of the teachers surveyed indicated that it is important for children to understand the equal sign (Question 5), but they gave a variety of reasons for why this is so. Two teachers stated that understanding the equal sign is necessary for understanding operations (specifically addition and subtraction): "Yes – just because… is that an answer? Well…it helps them understand + and -." Three teachers actually gave operator reasons, such as "because the equal sign leads to the answer." Six teachers stated that it is important to understand the equal sign because it occurs so often in math and because it is just "part of the math sentence." Surprisingly, just one of the 17 teachers surveyed stated that understanding the equal sign is important because students must learn about

Figure 5-1. Categories of responses to first question: Would you say your students have a fairly good understanding of what the equal sign means? ($N = 17$).



balancing sides of an equation and reasoning about equivalence. Five teachers either gave no response or responses too vague to categorize appropriately. If teachers do not realize the importance of the equal sign in both lower and higher mathematics, then they cannot be expected to instruct and review a relational interpretation of the equal sign in mathematical contexts. Unfortunately, responses from the surveyed teachers indicate that many teachers may not understand the role of the equal sign in mathematical contexts, and therefore may not be disseminating a relational view to their students.

Instructional Approaches and Teachers' Suggestions

Because elementary students so often fail tests of equivalence and give operator definitions of the equal sign, it is important to find out what teachers are, or are not, doing in the classroom to instruct children about the equal sign and other mathematical symbols. Determining what teachers are already doing is essential for deciding what else teachers can do or what should be changed about their approaches so that students learn to map their understanding of equivalence onto the equal sign. Teachers were asked in Question 2 whether they taught their students about the equal sign or whether students had entered their classes already knowing what the symbol meant. Five of the teachers (from Groups 2 and 3) indicated that their students already knew what the symbol meant so they did not have to teach their students about it. The remaining 12 teachers responded by saying they needed to teach it (six), they had to review the symbol's meaning (two), they had to teach just some of their students (one), or their students had an operator understanding of the equal sign that needed to be addressed in class (three). Because such a high proportion of teachers indicated they needed to instruct their students about the equal sign, it was fortunate several other questions on the survey were directed toward teaching methods.

With the third question I asked teachers to indicate how they had taught or reminded children about the meaning of the equal sign (or how they would if the need arose). Responses were categorized around the following themes: use of manipulatives or visuals only (e.g., balance beam, teeter-totters), mathematics contexts only but with various mediums (e.g., adding and subtracting money), symbolic arithmetic contexts only (e.g., addition problems with the "blank lines" in different places throughout the equations), verbal reminders and discussions (e.g., the equal sign means both sides are "fair"), a variety of approaches (i.e., use of manipulatives, verbal discussions and mathematics contexts), or other (i.e., teachers gave vague responses, such as "pointing to examples," that could not categorized any other way). As evident in Figure 5-2, many teachers (35.3%) stated that they would (or had used) a variety of approaches to teach children about the equal sign.

One of the teachers who responded with arithmetic contexts only stated that she would instruct students that the equal sign means putting the solution at the end of the problem: "I would explain to them that the equal sign represents what the numbers add up (or subtracted) to and that answer goes after the sign." This particular teacher, aside from suggesting that she would impart just an operator view of the equal sign, also indicated on her questionnaire that perhaps she does not fully understand the meaning of the equal sign: "I don't know if I really know what the equal sign is. If they understand that an answer comes after it, I think that is sufficient." Teachers who do not view the equal sign in a relational manner cannot be expected to impart a relational interpretation on to their students. Kieran's (1981) suggestion that an adult still may hold an operator view of this symbol may apply, unfortunately, to some teachers.

The fourth question was specifically geared toward determining whether teachers used a variety of contexts to display and contemplate the equal sign in their classrooms and, if so, what those different contexts looked like. Although just 35.3% of the teachers reported using a variety of contexts and approaches to *teach* students about the equal sign, 64.7% reported *using* multiple contexts in their classrooms. Teachers provided a list of contexts and examples they used in the classrooms that were organized into the following categories including: primarily arithmetic contexts (e.g., arithmetic problems on flash cards and in word problems), just one other context (e.g., money contexts, such as comparing values of different coins), various mathematics contexts (i.e., fractions and decimals), and multiple contexts (Figure 5-3). More than half of the teachers (58.8%) explicitly mentioned using money contexts. Despite so many teachers indicating that they use money contexts in their classrooms, the relatively low success rate (less than 50%) on the third clinical interview question, which required children to draw an equal sign between two nickels and one dime, suggests that children may not be recalling or generalizing their money equivalence experiences adequately. Although children may know that two nickels "equals" the same amount of money as one dime, it may require more instruction or deliberate training before they consistently and spontaneously represent this relation with the symbol of equivalence.

Figure 5-2. Categories of responses to the third question: Suppose you had to teach or remind your students about what the equal sign means. How would you go about doing that? ($N = 17$).

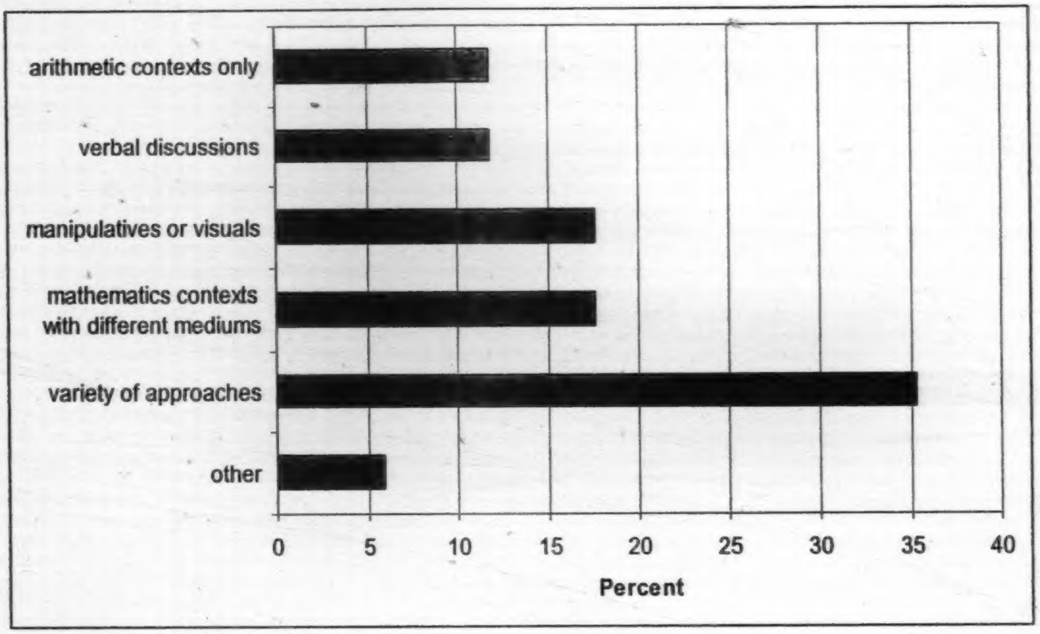
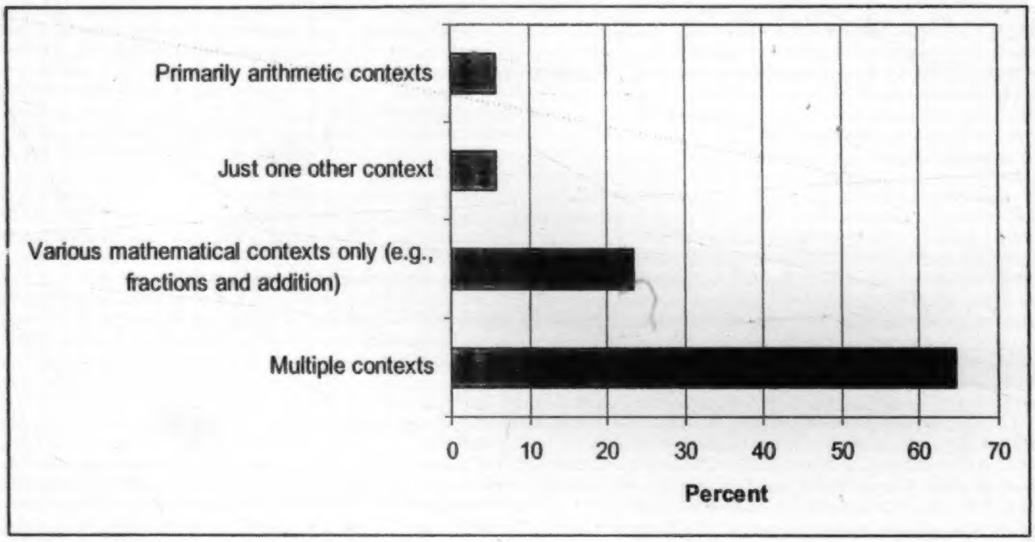


Figure 5-3. Categories of responses to the fourth question: Have you used the equal sign in different contexts with your student (such as showing equal amounts of money, e.g., 5 pennies = 1 nickel), or primarily in the context of arithmetic questions? ($N = 17$).



In the sixth question, I asked teachers to suggest ways to introduce children to the equal sign, resulting in comments similar to those given for Questions 3 and 4. Approximately 75% of the teachers suggested using manipulatives or a variety of approaches for introducing young students to the equal sign.

Teachers' Estimates About Success on Equivalence Problems

A sense exists among researchers that teachers are unaware of the extent to which children fail equivalence problems. Few researchers have investigated whether teachers are indeed unaware of children's misunderstanding of equivalence in mathematical contexts, however. Because so many arithmetic problems presented in textbooks and worksheets are canonical, it is very likely that teachers may have a great idea about which students are quick and accurate solving from left to right, but they may not realize that students are not thinking about equivalence on both sides of the equal sign. Rather, children with a strong addition schema can very efficiently solve from left to right without even considering equivalence. On the seventh question, teachers were asked to estimate how many children from each grade (Grades 1-5) they believed would correctly solve 3 + 4 + 2 = 1 + __. Eleven teachers provided estimates for each grade, three estimated for one or several grades (typically just the grades they instructed), two did not estimate at all, and one teacher estimated but qualified her best guesses by saying that accuracy would depend on the time of the year and the ability level of the students. In general, teachers estimated lowest accuracy for children in the lowest grade, with improved accuracy for each successive grade. The average numbers of children estimated to correctly solve the equivalence problem were 32.31% (Grade 1), 50.83% (Grade 2), 62.73% (Grade 3), 78.64% (Grade 4), and 91% (Grade 5).

What immediately stands out about these estimates is that they are grossly above the number of children who typically solve equivalence problems correctly. For example, just twelve of the fifty-seven Grade 2 children (21%) who received the screening test in the microgenetic study (Chapter IV) solved one or more equivalence problems correctly, compared to the estimate of 50% provided by teachers. Other researchers have found even lower performance, such as the remarkably low proportion (approximately 10%) of fourth- and fifth-grade students who successfully solve such problems (Perry, 1985, as cited in Perry, Church & Goldin-Meadow, 1988). One Grade 1 teacher in this survey estimated that 40% of her students would solve the equivalence problem correctly. After estimating, she actually administered the problem to her students:

> I was overly optimistic! I gave them the question on the board *after* estimating that 40% would "get it." 0% got it correct! They added 3 + 4 + 2 + 1 = **10**. They *ignored* the equal sign! So much for thinking that they know what the sign means!

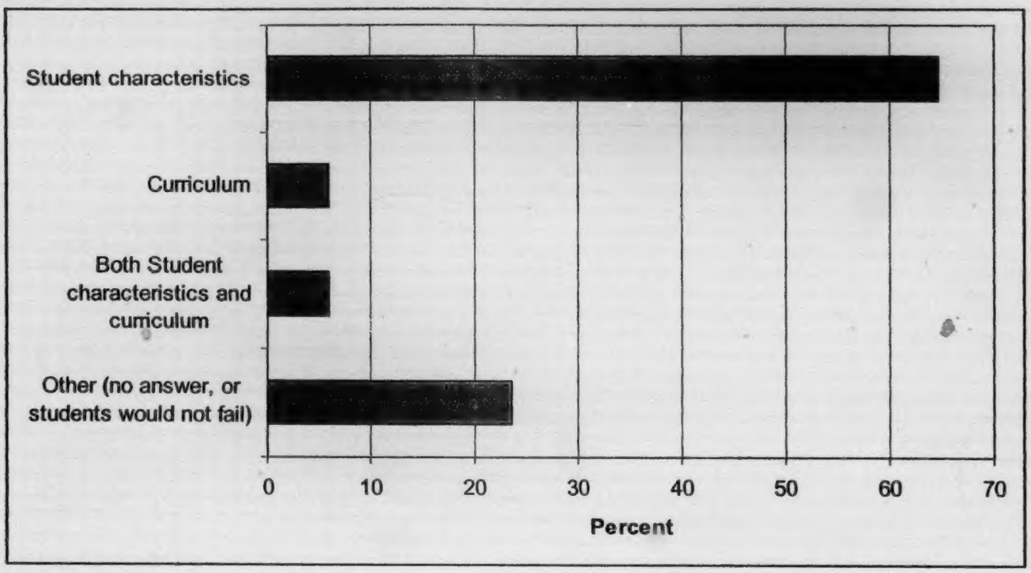
Based on the estimates found in this survey, it appears as though elementary school teachers may be overestimating children's ability to solve equivalence problems, in addition to overestimating how many students understand the equal sign in a relational manner. If teachers are unaware of the extent to which their students fail equivalence problems, then they are unlikely to target instruction to helping children overcome failure and address underlying misunderstanding of equivalence.

Teachers were then asked to indicate why they think some of their students might have failed to solve the equivalence problem (3 + 4 + 2 = 1 + __) correctly. Responses were coded in a similar manner as that done by Asquith et al. (2006) to see whether teachers' ideas about why children might not succeed on the equivalence problem in the present survey would be similar to those in the previous study. Teachers' responses were categorized to reflect one of three categories used by Asquith et al.: student characteristics, such as motivational factors, cognitive ability, and experiences with notation; curriculum, such as not enough focus on a particular concept or symbol; or teachers' own classroom observations and instructional practices, including abstaining from teaching a particular concept because they believed students had covered the topic in a previous grade. As evident in Figure 5-4, most teachers (64.7%) attributed children's difficulty with the equivalence problem to student characteristics such as difficulty with arithmetic or not trying hard enough. Asquith et al. found an almost identical proportion of teachers (63.6%) who attributed students' difficulty on the equivalent equations task to student factors such as math ability and motivational levels. None of the teachers surveyed in this study attributed difficulty on equivalence problems to the math curriculum, and neither did any of the teachers in the Asquith et al. study. Almost a quarter of the teachers I surveyed (23.5%) gave responses that could not be categorized using Asquith et al.'s scheme. For instance, some teachers simply stated that none of their students would fail the problem, therefore they did not have any guesses as to why some other students would get the problem incorrect.

On the ninth question, teachers were asked to state the strategies they believed their students would use to solve an equivalence problem such as 3 + 4 + 2 = 1 + __. Of the 15 teachers who answered this question, eight described various arithmetic procedures, such as adding all terms to the left of the equal sign and then subtracting the addend(s) to the right of the equal sign to calculate the solution. Three teachers believed that their students would use manipulatives to help them solve the problem, and two teachers thought students would use a variety of techniques, including the use of manipulatives, to solve the equivalence problems. One teacher indicated that her students would likely use an add-all approach, adding all the terms together rather than ensuring both sides of the equation were equivalent. Another teacher indicated she did not know how her students would approach an equivalence problem. In fact, many students use add-all approaches to solve equivalence problems, even though just one teacher responded with this strategy. Thus, in addition to overestimating accuracy on such problems, teachers may not have an accurate view of the types of strategies students use to solve equivalence problems.

On the last survey question, teachers were asked to estimate the number correct (out of ten) their students would demonstrate on an equivalence test. Ten equivalence problems, similar to those used in Chapters II-IV, were presented and teachers were asked to guess their students' average score. Among the four teachers of the youngest grades who gave estimates (Group 1), two estimated their students would solve seven of the ten problems correctly, one thought their students would only get two problems correct, and another teacher thought her students would solve four problems correctly, for a mean overall estimate of 50%. The teacher who estimated that her Grade 1 students would solve 40% of the problems correctly then administered the ten equivalence problems to her students:

Figure 5-4. Categories of responses to the eighth question: If you think that some students in your class would fail, why do you think they would fail? (*N* = 17).



> Again, I was overly confident! I have a below average class of 17 with four special needs. After estimating that the average score would be 4/10 I gave them a chance to do it. Three got 4/10, and 14 got 0/10. A few would not even attempt it.

Three teachers from Group 2 provided estimates, with one teacher believing her students would get all problems wrong, and the other two teachers thought their students would solve 8 of the 10 problems correctly. Among the 57 Grade 2 students who received the screening test for the microgenetic study (Chapter IV), the average score was 12.6%, considerably lower than what two of the Group 2 teachers estimated.

All seven teachers from Group 3 thought their students' average scores would be 7 of 10 or higher, with a mean score of 8.6. In fact, most studies report drastically lower accuracy by elementary school children than what teachers surveyed in this study estimated. Two of the Grade 2 teachers who completed the surveys had students who participated in the microgenetic study (Chapter IV), but unfortunately both tended to overestimate the success their students would have on the equivalence problems compared to the actual accuracy levels from the previous study. For instance, one of the Grade 1-2 teachers believed that her students would have 70% accuracy on the ten-problem test of equivalence, and the other thought that 70% of her students would solve the single equivalence problem correctly. Both estimates were significantly above actual student performance by students in both classrooms.

General Discussion

Little is known about teachers' views of students' understanding of the equal sign, how teachers actually instruct students about the equal sign in the classroom, or how well teachers think students will perform on equivalence problems. A relation between teachers' perceptions and children's performance does exist, however, so discovering what teachers' perceive about students' equivalence understanding is very important for determining how to correct and prevent problems (Asquith et al., 2006; Nathan & Koedinger, 2000a). Ten survey questions were posed to address three main questions: (a) how do teachers perceive students understanding of the equal sign, (b) how do teachers instruct students about the equal sign, and (c) how well do teachers think students will perform on tests of equivalence? Answers to the three main questions are discussed below.

First, most of the teachers believe their students understand what the equal sign means, even teachers with the youngest students. This result contrasts sharply with data from students of various grade levels (i.e., Grades 1-5), almost all of whom give operator, or incomplete, definitions of the equal sign (Chapters II & III). Teachers may be unaware of students' misunderstanding for one of two reasons: (a) children's relative success on canonical problems, which are used often in tests of arithmetic skill and speed, as well as in standard mathematical textbooks (Seo & Ginsburg, 2003), and (b) students often give relational interpretations of the equal sign when asked what it means in non-arithmetic contexts, such as comparing different coloured Cuisenaire rods (Seo & Ginsburg, 2003). Unfortunately, relational definitions of the equal sign in non-arithmetic contexts are not necessarily predictive of relational definitions in arithmetic contexts. If teachers are unaware of children's misunderstanding, they are unlikely spend enough time and resources teaching and reviewing a relational interpretation of the symbol, particularly in arithmetic contexts.

All teachers stated that it is important for children to understand the equal sign, but only one of the 17 teachers stated that the equal sign is important for representing balancing two sides of an equation. Other teachers reported that the equal sign is important for understanding operations (i.e., + and -), for pointing to the answer, or because it is part of the math sentence. If teachers themselves do not realize the importance of the equal sign in mathematics then they cannot be expected to instruct and review a relational interpretation of the equal sign in mathematical contexts. Kieran (1981) suggested that even some adults may not have a relational view of the equal sign in mathematics, a conclusion supported by at least one teacher in this survey who admitted that she only taught students an operator understanding and that perhaps she did not know what the equal sign really means. I suspect that most teachers do hold a relational interpretation of the equal sign, but because of their proficiency with mathematics, particularly the mathematics taught at the lower grade levels, they overlook the importance of deliberately teaching children about the relational role of the symbol. The limited or biased instruction resulting from teachers' advanced knowledge in a particular domain, in this case the equal sign in mathematics, has been supported by other researchers, and is referred to as the "expert blind spot" by Nathan and Petrosino (2003). Results from the first two questions on the survey suggest that teachers may not realize the extent to which children hold misconceptions about the equal sign and that few of them explicitly recognize the importance of the symbol in mathematics.

Second, many teachers reported using a variety of approaches to discuss equivalence and the equal sign with their students, and they suggested the use of multiple approaches to teach children who still need to learn about the symbol. Several hands-on approaches were mentioned multiple times by teachers of various grade levels. In particular, teachers reporting using the equal sign in the context of money (e.g., 1 dime = 10 pennies), social situations (e.g., there are an "equal" number of boys and girls in each group), and balance scales or teeter-totters. Two Grade 1 teachers indicated that they incorporate the equal sign into a variety of activities every day. For example, one teacher stated: "I have the concept of equals incorporated into many of our classroom routines…such as the special person of the day…their name contains two vowels plus three consonants, which *equals x* letters in all." Another teacher described a very deliberate attempt to integrate discussion of equality into everyday discussion:

> I start by using words like "fair" (e.g., two candy worms versus three candy worms…is this "fair?"). Then I start saying "equals", "same as", and then more consistently I use "equals". When I ask the students what "equals" means, they say "fair."

Although teaching children about equivalence with a variety of mediums and contexts may seem like a responsible, best-practices approach to equivalence instruction, children's failure on equivalence problems and misconceptions of the equal sign indicate that this approach (i.e., using multiple strategies and contexts) may not be sufficient. Children must be able to link their concept of equivalence to the symbol and use their understanding of the equal sign to solve arithmetic and algebra problems, not just balance a scale and count to determine that the number of boys and girls on each team is "equal." Although research on potential instructional techniques and interventions is still ongoing, several implications have emerged. First, children may be best served by experiencing arithmetic problems in noncanonical, not just canonical, arithmetic formats so they must learn to reason about arithmetic on both sides of the equal sign, not simply solving from left to right and mindlessly putting the sum after the equal sign (Sáenz-Ludlow & Walgamuth, 1998; Seo & Ginsburg, 2003). In the microgenetic study I found that after Grade 2 students solved numerous equivalence problems over several sessions, their adherence to the addition schema actually lessened, as evident by their improved operational patterns scores. Second, children who learn procedures to solve equivalence problems or who are taught about the principle of equivalence in arithmetic, rather than just talking about "fairness" or balance, do show some improvements on tests of equivalence (e.g., Perry, 1991). Third, children's ability to reiterate instructions for solving equivalence problems is enhanced if teachers use spontaneous gestures in their explanations of solving equivalence problems (Goldin-Meadow et al., 1999).

The three examples of emerging instructional recommendations all center around one theme: focusing discussions, examples, procedures, explanations, and gestures around the equal sign in arithmetic, particularly noncanonical situations such as equivalence problems. The focus implied by teachers surveyed in the present study is largely about teaching children to reason about equivalence in a very general sense, such as equivalence between coin combinations, candy, and balancing a teeter-totter. Children's difficulty with equivalence problems and interpreting the equal sign does not appear to stem from a misunderstanding of equivalence in a general sense, however. Mix (1999) demonstrated that even preschool children can recognize equivalence across numerical sets, even sets that differ in domain, such as visual and audio domains. Other researchers have commented that, from a very early age, children discuss equality (e.g., "your cookie is the same size as my cookie"), and certainly recognize situations of inequality (Seo & Ginsburg, 2003). In fact, some kindergarten children can even solve equivalence problems when they are presented with manipulatives (Study 1, Chapter II). The trouble really seems to be helping children map their general understanding of equivalence to the symbol of equivalence in mathematical contexts, not thinking about equivalence in general.

Teachers who are unaware of the extent to which children hold misconceptions about the equal sign may be focusing on skill sets that children already possess (i.e., reasoning about equivalence, broadly defined), rather than focusing on the symbol that is overwhelmingly misunderstood. Nathan and Koedinger (2000a, 2000b) described this phenomenon of attributing poor performance on specific mathematic skills to general misconceptions rather than difficulty with symbols, referring to it as the symbol-precedence view. The teachers surveyed in this study undoubtedly have their students' best interests at heart, and are likely investing a great amount of time and energy into ensuring that children learn to think about equality in mathematics *and* social contexts. An important next step now appears to be helping teachers understand their symbol-precedence view and students' misconceptions of the equal sign in particular, so that teachers can then channel their creativity and instruction into helping children map a relational understanding of the equal sign to problems in mathematical contexts.

Third, many teachers overestimated children's ability to correctly solve arithmetic equivalence problems. This overestimation is not surprising considering many also reported (incorrectly) that their students fully understand the equal sign. Three of the teachers surveyed had students who actually solved the equivalence problems, allowing me to directly compare students' performance with teachers' estimations. Two of the teachers had students in the microgenetic study (Chapter IV) and the other teacher administered the equivalence problems from the survey to her students and reported on their performance. In all three cases, teachers greatly overestimated both how many children would get the sample equivalence problem correct and the average score their classroom would receive on a test of equivalence problems. If teachers are unaware that children have difficulty solving these challenging arithmetic problems, then students are unlikely to benefit from direct instruction, either procedural (i.e., add up all terms on one side of the equal sign and subtract the term(s) from the other side) or principle-based (i.e., both sides of the equal sign must have the same total value) in nature. When asked to consider the strategies children might use to solve an equivalence problem, just one teacher reported an add-all approach, which is an extremely common approach for equivalence problems, particularly when the blank line occurs at the very end of the equation (e.g., a + b + c = a + __) rather than immediately after the equal sign (e.g., a + b + c = __ + c) (McNeil & Alibali, 2004). Thus, in addition to overestimating accuracy, teachers may not have a clear idea of the type of strategies children use to solve equivalence problems.

In summary, teachers appear to be using a variety of approaches to instruct students about equivalence and to bring discussions of "equals" into daily conversation, yet they overestimate students' ability to solve equivalence problems and are not aware that students misunderstand the equal sign in arithmetic contexts. Teachers' time and energy might be better spent on exposing children to noncanonical problems, such as equivalence problems, and on focusing on instilling in students a relational understanding of the equal sign in the context of math. Greeno and Hall (1997) suggested that representations, including symbols such as the equal sign, should be taught as tools, not as ends in themselves. Children cannot effectively use the symbol of equivalence if they do not understand it in a relational sense. If children simply learn the symbol as a symbol unto itself, such as the "thing that comes before the answer," rather than its role in the larger representational system of mathematics, then they are limited in their ability to correctly solve equivalence problems and face higher mathematical concepts such as algebra. Furthermore, Hiebert and Carpenter (1992) argued that the goal of teaching mathematics is to bridge the external (symbols) to concepts, such as understanding of adding, subtracting, and equivalence. Again, instructional techniques that support the mapping of concepts to the symbol system, rather than assuming that children have no difficulties with the symbols, may be the most effective way to help children overcome their misunderstanding of the equal sign in mathematical contexts.

Children's performance and conceptions cannot be expected to change and improve until teacher perceptions are effected. A possible intervention, in addition to the types of interventions presented in the literature so far, might simply be educating teachers about the topic and supporting modifications in classrooms to focus on the equal sign in mathematical contexts. Future research is needed to explore the effects of informing and educating teachers about children's misconceptions, particularly on children's views of the equal sign and performance on equivalence problems. The most practical and efficient intervention of all may be the simplest to administer: Simply let teachers in on the secret.

References

Asquith, P., Stephens, A. C., Knuth, E. J., & Alibali, M. W. (2006). *Middle school mathematics teachers' knowledge of students' understanding of core algebraic concepts: Equal sign and variable*. Manuscript submitted for publication.

Baroody, A. J., & Ginsburg, H.P. (1983). The effects of instruction on children's understanding of the "equals" sign. *The Elementary School Journal, 84*, 199-212.

English, L. D., & Halford, G. S. (1995). *Mathematics education: Models and processes*. Mahwah, NJ: Erlbaum and Associates.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teachers' hands tell the student's mid about math. Journal *of Educational Psychology, 91*, 720-730.

Greeno, J. G., & Hall, R. P. (1997). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan, 78*, 361-367.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: MacMillan Publishing.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*, 317-326.

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development, 76*, 883-899.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28*, 451-466.

Mix, K. S. (1999). Preschoolers' recognition of numerical equivalence: Sequential sets. *Journal of Experimental Child Psychology, 74*, 309-332.

Nathan, M. J., & Koedinger, K. R. (2000a). An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction, 18*, 209-237.

Nathan, M. J., & Koedinger, K. R. (2000b). Teachers' and researchers' beliefs about the development of algebraic reasoning. *Journal for Research in Mathematics Education, 31*, 168- 191.

Nathan, M. J., & Petrosino, A. (2003). Expert blind spot among preservice teachers. *American Educational Research Journal, 40*, 905-928.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359-400.

Sáenz-Ludlow, A., & Walgamuth, C. (1998). Third graders' interpretations of equality and the equal symbol. *Educational Studies in Mathematics, 35*, 153-187.

Seo, K., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence…": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stephens, A. C. (2006). Equivalence and relational thinking: Preservice elementary teachers' awareness of opportunities and misconceptions. *Journal of Mathematics Teacher Education, 9*, 249-278.

CHAPTER VI

GENERAL DISCUSSION AND CONCLUSIONS

Successfully and accurately interpreting symbols is vital for basic literacy, and in the domain of mathematics, simple relations and complex proofs alike are described using conventional symbols. Although elementary school children typically learn to read and use many symbols, including Arabic numerals, relatively quickly, at least one mathematical symbol is consistently misunderstood: the equal sign. The pivotal role of the equal sign in higher mathematics such as algebra, combined with the fact that many teachers are unaware of children's misunderstanding, makes the examination of potential interventions an interesting and important topic of research. In a series of five studies, I examined children's views of the equal sign, performance on equivalence problems, and adherence to operational patterns in relation to specific arithmetic experiences (e.g., exposure to nonsymbolic problems first) and in the broader context of students' instructional experiences as described by elementary school teachers.

The five studies were useful for detecting the effects of specific arithmetic experiences and instruction on children's ability to solve equivalence problems and their views of the equal sign and operational patterns. In the first three studies, I examined whether children enter school already having an understanding of arithmetic equivalence, whether nonsymbolic and semi-symbolic experiences are beneficial for mapping relational strategies to symbolic equivalence problems, and the effects of minimal symbolism on performance (i.e., the semi-symbolic condition). In the fourth study, I explored changes in children's performance on tests of operational patterns, and I also tracked individual learning paths and the five elements of change identified by Siegler

(1996) among three different exposure conditions for Grade 2 students. Finally, in the fifth study I examined teachers' perceptions about children's (mis)understanding of the equal sign and their instructional approaches for equivalence and the equal sign.

To briefly recap what is more thoroughly described in the general discussions of each chapter, in Study 1 I found that indeed some children do enter school already possessing the ability to use arithmetic to make both sides of an equation quantitatively equal, at least when solving in an unconventional, nonsymbolic condition. In fact, children solved more problems correctly than would be expected if they had guessed, and some solved all ten problems correctly. Unfortunately, the early invention of a relational strategy for solving equivalence problems relationally often does not get mapped onto the symbol for equivalence, as evident by elementary school children's low accuracy on equivalence problems throughout numerous studies, and by their operator definitions of the equal sign. Nonetheless, kindergarten children's ability to invent strategies for solving nonsymbolic equivalence problems, whether relational or operator, is remarkable, considering (a) these children have received little, if any, formal mathematical instruction, and (b) the relative complexity of these problems compared to problems kindergarten to which children typically would be exposed.

Kindergarten children varied widely with respect to the number of equivalence problems they solved correctly and the verbal justifications they gave to support their solution options. Children with either very high (i.e., at least five out of ten problems correct) or extremely low (i.e., no problems correct) accuracy demonstrated high rates of agreement between their solution options and their verbal justifications. Children in these two groups demonstrated convincing evidence of using a particular strategy, with most children with 0% accuracy inventing an add-all or add-to-equal strategy. Children with low (but not zero) accuracy, in contrast, demonstrated less agreement between their solution options and verbal justification. Low agreement scores might be attributed to the fact that children could not verbally explain why and how they had chosen to solve the problems, or they might reflect children's attempts to improve their chance of using an accurate strategy. That is, by doing one thing and saying another, children double their chance of appearing accurate, as opposed to matching their verbal justifications to their solution option.

I expanded on previous research with Grade 2 students (Sherman & Bisanz, 2004) by using a within subject-design and a multiple-choice paradigm. This methodological adaptation led to the demonstration, for the first time, that Grade 2 children's performance on symbolic equivalence problems can improve after experiencing equivalence problems in either a nonsymbolic or semi-symbolic context (Chapters II, Study 2, and Chapter III). For these students, solving equivalence problems in the absence of all symbols, or with just some conventional symbols, is enough to maintain high accuracy on symbolic problems presented a week later. In contrast, receiving symbolic problems first can have differential effects on subsequent performance, depending on the condition equivalence problems are presented in. For instance, when children solved nonsymbolic problems in their second session, accuracy improved dramatically, indicating that they approached nonsymbolic problems very differently than they had approached the symbolic problems from the first session. However, when children solved semi-symbolic problems in their second session, accuracy remained very low. Perhaps the visual similarities between the two conditions were sufficient for children to apply their faulty, operator-based addition schemas, resulting from prior experience with symbolic problems. More research is needed to positively identify the number of symbols, or specific combination of symbols, necessary to evoke children's addition schemas, as well as delineate the circumstances in which these symbols elicit children's schemas. For instance, is the equal sign alone enough to evoke the schema, are other operational symbols (e.g., "+") necessary, and are schemas only evoked if they have been previously activated (i.e., a previous symbolic session), and if so, how close to the two encounters need to be to for children to be influenced by prior experience? In Chapter III, children who began with semi-symbolic problems had very high accuracy in their first session, indicating they did not apply an addition schema based on that particular combination of just some conventional symbols.

Children's changing views of the equal sign, performance on equivalence problems, and scores on tests of operational patterns were explored over numerous sessions using a microgenetic design. Based on the studies in Chapters II (Study 2) and III, I hypothesized that children who received nonsymbolic or semi-symbolic exposure phases, consisting of six equivalence problems at beginning of each session, would require fewer sessions to consistently and accurately solve symbolic equivalence problems compared to children who only ever experienced symbolic problems. There was no difference between Grade 2 children in the three exposure condition groups, however, prompting me to explore other sources of change, such as the onset of feedback during exposure phases. The lack of a difference also prompted me to hypothesize potential boundary conditions surrounding the effectiveness of nonsymbolic and semi-symbolic conditions as interventions for improving accuracy on symbolic equivalence problems. Performance on symbolic equivalence problems may be modified by how accurately children solve exposure problems, and accuracy on symbolic problems may be directly influenced by receiving feedback on exposure problems. More than half of the children who received feedback at some point during the equivalence sessions made dramatic improvements (i.e., 75%) on symbolic equivalence problems within one or two sessions.

The microgenetic study was also instrumental in highlighting the different paths of change that Grade 2 children displayed. That is, some children experienced immediate success on symbolic problems, meeting the stringent stop criterion after just three or four sessions, whereas other children required numerous sessions to show improvements. Five children served as examples to demonstrate the five paths of change, as they highlighted the timing and rate of change in accuracy, how confident children were, and the number of strategies used across the sessions. In general, children who learned to correctly solve symbolic equivalence problems correctly and quickly (i.e., in three to five sessions), were better able to maintain their relational approaches and transfer their skills to novel problems presented two weeks later compared to children who required more sessions (i.e., six to nine). This study was the first to show that children's ability to solve and reconstruct symbolic equivalence problems could improve after numerous equivalence sessions alone, as opposed to direct instruction about operational patterns and the equal sign, with overall accuracy improving across all subtests as quantified and scored for this particular study. Children's ability to reconstruct equivalence problems after viewing them for five seconds appears to be a particularly useful task for predicting which children might learn to solve symbolic equivalence problems most quickly.

Finally, in the last study, I examined teachers' views about children's understanding of the equal sign, as well as how teachers described their approaches to instructing students about equivalence and the equal sign in mathematics. The teachers I questioned appeared to be largely unaware of the extent of children's misunderstanding, often grossly overestimating children's success on symbolic equivalence problems compared to performance from previous research (including Chapters II and III). The experiences teachers described providing in the classroom to foster a relational view typically occurred in canonical formats, with an operation to the left and the result to the immediate right of the equal sign. Unfortunately, use of canonical arithmetic examples may be unintentionally promoting an operator interpretation, much like the situation found by Seo and Ginsburg (2003). The impressive range of experiences teachers provide for their students to learn about equivalence in general may be more effective in promoting a relational view of the equal sign if they specifically discuss the role of the equal sign in arithmetic, provide noncanonical examples of arithmetic equations, and rehearse equivalence problems. More research is needed, however, to find the most effective intervention for the classroom. The study represents a first step to learning about typical classroom practices and teacher perceptions.

Although numerous questions were addressed with the five studies, new questions have been raised as a result, and there remain many opportunities for future research. For example, I would like to pursue research on children's adherence to the addition schema and how adherence relates to equivalence problems presented in even more highly modified conditions. Specifically, I would like to identify the number or combination of mathematical symbols that is necessary and sufficient to activate children's addition schema more concisely than was possible using just the semi-symbolic condition (such as by using problems presented where the *only* symbol is the equal sign). I believe the relation between performance on equivalence problems and adherence to the addition schema has important implications regarding children's mental representation and the factors that influence those representations (Dixon, 2005). Exploring this question would help answer whether children's difficulty with equivalence problems relates primarily or only to the equal sign, and would have subsequent instructional implications.

Another area I wish to explore is the relation between children's difficulty on equivalence problems and their cognitive inhibition skills. I suspect that children who have great ability to inhibit prepotent responses may be better able to suppress their addition schemas when solving symbolic equivalence problems. Being able to identify children who may be more susceptible to problems on equivalence problems (i.e., those with poor inhibition skills) may be helpful for tailoring deliberate equivalence and equal sign instruction. In future studies I would like to assess potentially relevant factors such as language and general mathematics skills that will allow me to better describe my samples and test for important relations.

The results from the microgenetic study have raised many new questions related to individual differences, variability in strategies within children, the robustness of the change resistance account, and perhaps most surprisingly, potential limitations surrounding the efficacy of nonsymbolic or semi-symbolic conditions as interventions techniques. Hypothesized benefits for children in the nonsymbolic and semi-symbolic conditions did not result from the current design in some analyses, raising additional questions about the boundary conditions for the effectiveness of certain interventions.

Ensuring success on nonsymbolic and semi-symbolic problems, such as providing feedback and tailored instructions, may be vital for enabling children to map their strategies onto similar problems presented with conventional symbols.

Despite the number of questions remaining, just some of which are listed above, my most immediate interest is in the realm of interventions. Researchers studying children's misunderstanding of the equal sign have constructed a convincing body of research suggesting that elementary children have extremely poor performance on equivalence problems that is often difficult to change, and they possess a stubbornly resistant operator interpretation of the equal sign in arithmetic contexts. Recently researchers have found that teachers appear unaware of children's misunderstanding and dismal equivalence performance, including my own research presented in Chapter V, leading me to question what type of interventions would be most effective, efficient, and convincing to educators, and to wonder about the process for deciding upon interventions when the need for one is recognized. We now have research support for changes in instructing children about the equal sign and practice with noncanonical arithmetic, so the time might be right for examining how educators, textbook editors, and schools or education navigate the process of curriculum modification. Furthermore, little is known about what the best intervention for alleviating children's misunderstanding of the equal sign would be. Educating elementary teachers about this misconception may be the most efficient, direct way for producing effective change. Another option is changing the ways textbooks present the sign of equivalence and canonical problems. Perhaps a combination approach is necessary for ensuring that children achieve a relational view of the equal sign in mathematics.

I expect my research to draw from and contribute to various domains, including mathematical cognition, mental representation, children's understanding of symbols, mapping (e.g., linking strategies from nonsymbolic to symbolic equivalence problems), and the effectiveness of specific instructional techniques. Improving children's concepts of specific mathematical symbols, such as the equal sign, will likely involve additional empirical research as well as dialogue with educators, more thorough examination of current curriculum and textbooks, and some analyses of practical classroom interventions that can be instituted efficiently in all elementary schools. Children's (mis)understanding of the equal sign may appear to be a very specific problem, but its application extends to all situations in which children fail to map a general, important, concept onto its symbolic representation. The prerequisite understanding for symbolic skill is realizing the relation between a symbol and its referent (McMullen & Darling, 1996). In the case of the equal sign (and other symbols, such as the use of various letters to represent variables in algebra), children do not appear to be mapping the referent (equivalence) to the symbol.

Ultimately, being able to communicate, solve problems, and mentally represent ideas, depends on the ability to recognize, use, and manipulate conventional symbol systems. "Understanding how children acquire and use formal representational tools is particularly important at a time when the information revolution has the potential to provide children with new means of organizing and representing scientific information" (Miller, 2000, p. 24). If a chain is only as strong as its weakest link, then children's misconceptions of the equal sign, or any other symbol for that matter, cannot be ignored or trivialized. The next steps seem to be educating teachers about children's pervasive misunderstanding and determining the most effective way to prevent and correct children's misconception about the equal sign.

References

Dixon, J. A. (2005). Mathematical problem solving: The roles of exemplar, schema, and relational representations. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 379-411). New York: Psychology Press.

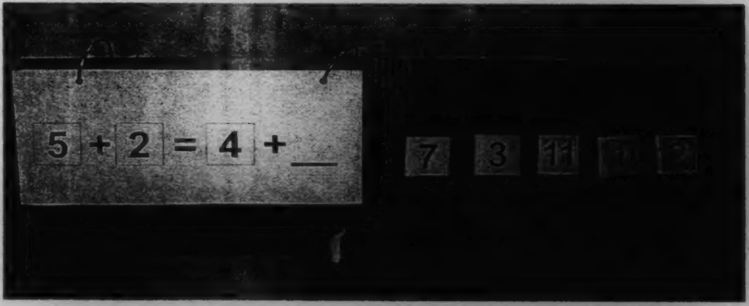
McMullen, N. B., & Darling, C. A. (1996). Symbolic problem solving: An important piece of the emergent literacy puzzle. *Early Child Developmental and Care, 121*, 25-35.

Miller, K. (2000). Representational tools and conceptual change: The young scientist's tool kit. *Journal of Applied Developmental Psychology, 21*, 21-25.

Seo, K., & Ginsburg, H. P. (2003). "You've got to carefully read the math sentence…": Classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Appendix A

Nonsymbolic	
Symbolic	
Semi-symbolic	

Appendix B

Classification	Criteria	Example Justification
Relational	Describes making both sides have the same numerical value.	"There's 7 on that side (left), so I put 2 with the 5 on this side (right) to make 7. There's 7 and 7 now."
Add-All	Describes adding up all terms and placing that value in the empty bin (or on the line); States there's as many in bin (or on the line) as all the others together.	"I put 12 because there are 12 all together."
Add-to-Equal	Describes adding up the terms to the left of the blue tent (or equal sign) and matching that value as the solution.	"There's 7 over there (left), so I put 7 here (last bin or on the line)."
Small Number	Describes choosing the smallest value.	
Question Mark (?)	Describes choosing the question mark because none of the other responses are correct.	"None of these other numbers are the right answer." *note: when children gave this response, the experimenter asked the child what number they were looking for to determine whether another approach was being used but the child had simply miscounted.
Pattern/Matching	Describes making or completing a pattern; Describes matching one of the existing terms.	"There's a 3 then a 2 on this side (left) and a 3 on this side (right), so we needed a 2 next." *note: often children would explicitly state that there was a pattern.
Ambiguous	Justification cannot be classified.	"no 4," without clear reasoning for why a 4 would be needed (no elaboration even after prompting).

Appendix C

	Part-Whole	Combination
Kindergarten	1 + 4 = 2 + __	3 + 1 + 2 = 3 + __
	4 + 3 = 2 + __	3 + 2 + 2 = 3 + __
	3 + 1 = 2 + __	2 + 3 + 2 = 2 + __
	5 + 2 = 4 + __	2 + 1 + 2 = 2 + __
	4 + 2 = 3 + __	2 + 4 + 1 = 2 + __
Grade 2	5 + 3 = 2 + __	2 + 1 + 2 = 2 + __
	6 + 2 = 4 + __	4 + 2 + 1 = 4 + __
	1 + 4 = 2 + __	2 + 4 + 4 = 2 + __
	5 + 2 = 4 + __	4 + 3 + 1 = 4 + __
	3 + 1 = 2 + __	2 + 4 + 1 = 2 + __
	4 + 5 = 3 + __	3 + 1 + 2 = 3 + __
	3 + 4 = 5 + __	5 + 1 + 2 = 5 + __
	2 + 4 = 3 + __	3 + 2 + 2 = 3 + __

Appendix D

Classification	Example Response
Relational	"It means that both sides need to be the same"
Operator	"It means the answer comes next" "You need to put the answer" "What the answer of the question is" "You need to add up the numbers and put the answer on the line"
Ambiguous	"It means equals"
I don't know	"I don't know what it means"

Appendix E

Equivalence	1	5 + 3 = 2 + __
	2	4 + 2 + 1 = 4 + __
	3	2 + 4 + 1 = 2 + __
	4	6 + 2 = 4 + __
	5	2 + 4 + 4 = 2 + __
	6	1 + 4 = 2 + __
	7	5 + 2 = 4 + __
	8	4 + 3 + 1 = 4 + __
	9	3 + 1 = 2 + __
	10	5 + 1 + 2 = 5 + __
Canonical	1	4 + 2 = __
	2	2 + 5 + 1 = __
	3	5 + 4 = __
	4	1 + 9 = __
	5	3 + 1 + 2 = __

Appendix F

Equation Solving	3 + 4 + 2 = 3 + __	
	6 + 3 = 4 + __	
	5 + 2 + 3 = 2 + __	
Problem Structure	Test 1	Test 2
	7 + 4 + 5 = 7 + __	6 + 7 + 2 = 2 + __
	3 + 6 = 5 + __	7 + 3 = 5 + __
	4 + 1 + 6 = 6 + __	5 + 2 + 4 = 5 + __
Equal Sign Definition	5 + 3= + 3 = 2 + __	“What does this math symbol (=) mean?”

Appendix G

It's definitely wrong	It's probably wrong	It might be wrong	I don't know if it's right or wrong	It might be right	It's probably right	It's definitely right

Appendix H

Not so smart	Kind of smart	Very smart

Appendix I

SESSION 2

Exposure Problems

	Type	Problem	Solution	Justification	Rating	Correct
1	Com	3 + 2 + 5 = 3 + __				
2	PW	8 + 1 = 3 + __				
3	PW	1 + 12 = 4 + __				
4	Com	4 + 6 + 2 = 4 + __				
5	Com	1 + 2 + 2 = 1 + __				
6	PW	5 + 7 = 2 + __				

Part-Whole (PW): ________

Combination (Com): ________

Total: ________

Test Problems

	Type	Problem	Solution	Justification	Rating	Correct
1	Com	2 + 7 + 4 = 2 + __				
2	3PW	1 + 5 + 2 = 3 + __				
3	PW	4 + 2 = 1 + __				
4	PW	2 + 9 = 5 + __				
5	3PW	10 + 1 + 4 = 2 + __				
6	3PW	4 + 2 + 1 = 7 + __				
7	PW	7 + 4 = 3 + __				
8	Com	1 + 5 + 4 = 1 + __				
9	Com	3 + 9 + 1 = 3 + __				

Part-Whole (PW): ________

Combination (Com): ________

Three-term Part-Whole (3PW): ________

Total: ________

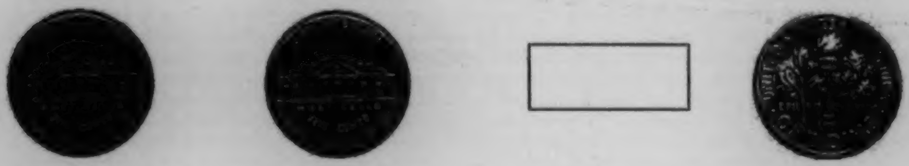
Appendix J

Clinical Interview Questions:

1. Here you see '2 + 4 = 3 + __.' Can you show me how you might write this problem without using numbers or other symbols?
2. How would you could you demonstrate the following using math symbols? If your math teacher asked you to turn the following into a math equation, how would you do it?

⋈ ⋈ ⋈ and ⋈ ⋈ ⋈ ⋈ ⋈ are the same as ⋈ ⋈ ⋈ ⋈ and ⋈ ⋈ ⋈ ⋈

3. If I wanted to show my friend that two nickels are the same amount of money as this one dime, how could I do that using just one symbol? What could I put in the empty box to show that two nickels are the same amount of money as one dime?



Appendix K

	Basic Hierarchical Classifications	Examples
	Use of pictures for numbers only	♥♥ ♥♥♥♥ ♥♥♥
	Use of written words for numbers only	Two Four Three
Problem 1	Insertion of operator symbols into picture or written representations	♥♥ + ♥♥♥♥ + ♥♥♥
	Written words for operator symbols	♥♥ plus ♥♥♥♥ plus ♥♥♥
	Representation for one operation	♥♥ ♥♥♥♥
	Correct representation (i.e., without any conventional math symbols) of the entire problem	♥♥ ♥♥♥♥ ⇄ ♥♥♥ □
	Use of numbers only	3 5 4 4
	Use of numbers and "+" for "and" only	3 + 5 4 + 4
	Use of numbers and "+", as well as the incorrect use of "+" for "are the same as"	3 + 5 + 4 + 4
Problem 2	Use of "=" at the end of the problem (incorrect position)	3 + 5 + 4 + 4 =
	Use of "=" before the last number (incorrect position)	3 + 5 + 4 = 4
	Use of "=" in the correct position.	3 + 5 = 4 + 4
	+	+
	10	10
	20	20
Problem 3	Other	⇄
	5 + 5 = 10	5 + 5 = 10
	=	=

Appendix L

Date: ________________________

Grade level mathematics you teach: ________

Dear Teacher:

Below are some questions I would like to discuss with you. Please feel free to look the questions over in advance and prepare responses if you would like to do so. I would like to remind you that you can refrain from answering any or all of the questions at your discretion, and you may terminate the discussion at any time. Responses will remain anonymous, although your input will be used to better understand performance by the students in your class and in other classes.

1. Would you say that your students have a fairly good understanding of what the equal sign (=) means? If so, how do you know? And if not, how do you know?

2. Did you teach your students what the equal sign means, or did they enter your class already knowing what it means?

3. Suppose you had to teach or remind your students about what the equal sign means. How would you go about doing that?

4. Have you used the equal sign in different contexts with your student (such as showing equal amounts of money, e.g., 5 pennies = 1 nickel), or primarily in the context of

arithmetic questions? Please provide some examples of the types of questions or situations in which your students would have seen the equal sign.

5. Do you think it is important that students understand what the equal sign is? If so, why? If not, why not?

6. How would you suggest that children be introduced to the equal sign?

7. Consider the following problem: 3 + 4 + 2 = 1 + __. Indicate on the following lines the percentage of children in each grade you believe would answer the problem correctly.

Grade 1: _____%
Grade 2: _____%
Grade 3: _____%
Grade 4: _____%
Grade 5: _____%

8. If you think that some students in your class would fail, why do you think they would fail? (If you teach mathematics at more than one grade level, please answer separately for each grade level)

9. How do you think students in your grade(s) generally solve this kind of problem? Can you describe their likely solution process, step by step?

10. Consider the following test questions. If your students were to complete a test with the following questions, what would you estimate their average score to be (out of ten)? Please indicate the number correct (out of ten) you would expect your students to get (on average).

5 + 3 = 2 + ___
4 + 2 + 1 = 4 + ___
2 + 4 + 1 = 2 + ___
6 + 2 = 4 + ___
2 + 4 + 4 = 2 + ___
1 + 4 = 2 + ___
5 + 2 = 4 + ___
3 + 2 = 1 + ___
4 + 3 +1 = 4 + ___
5 + 1 + 2 = 5 + ___

Your estimation for your students' average score (out of ten): ____________________